UNIVERSITY OF LONDON THESIS

Degree PhD Year 2007 Name of Author PATEL, Deven Dhirajkumar

Statistical modelling of virological and immunological patterns in HIV-1 infected pregnant women in Europe

A thesis presented for the degree of

Doctor of Philosophy

University of London

Deven Dhirajkumar Patel

Institute of Child Health

University College London

2007

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

I, Deven Dhirajkumar Patel, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

This thesis aims to investigate immunological and virological markers of HIV infection during pregnancy and to inform understanding of the impact of antenatal antiretroviral therapy (ART) regimens on such markers, using data from the European Collaborative Study (ECS). Since 1986, over 7,000 HIV infected pregnant women have been enrolled from 23 centres in 10 countries. Maternal information collected included timing and type of treatment regimen, repeated measurements on maternal CD4 counts and viral load (since 1987) and socio-demographic variables.

Repeated HIV RNA measurements over pregnancy in over 300 untreated women were examined using linear mixed effects (LME) models. Immunological (CD4) and virological changes over pregnancy in 162 women receiving highly-active ART (HAART) at conception and throughout pregnancy were examined with a piecewise LME model, using a conditional likelihood approach to account for left-censored measurements. A bivariate LME model was used to assess the correlation between the two markers. Viral response to initial HAART regimens in 240 ART-naïve pregnant women was determined using interval censored regression, with a propensity score to reduce treatment allocation bias. A Cox regression model was used to assess the effect of HAART on risk of premature delivery, treating elective caesarean section deliveries as right-censored outcomes.

HIV RNA viral load was estimated to decrease over the second and third trimesters of pregnancy in untreated and treated women. Differences in levels of HIV RNA viral load and CD4 count by race and history of injecting drug use were identified. The rate of achieving undetectable viral load was greater for women initiating on nevirapine-containing HAART than for women on protease inhibitor-based HAART. There was no evidence of an increased risk of birth defects or prematurity associated with type or timing of HAART. These findings contribute to the evidence-base for management and understanding of HIV infection in pregnancy.

Statistical modelling of virological and immunological patterns in HIV-1 infected pregnant women in Europe

Appendices

List of Tables

List of Figures

Acknowledgements

I would like to thank my supervisors, Marie-Louise Newell, Claire Thorne and Mario Cortina-Borja for their invaluable advice, support and encouragement over the years.

It would like to thank all the clinicians and laboratory staff who have contributed to this research through their collaboration in the European Collaborative Study. I would especially like to thank the parents and carers who participated in the survey.

I would like to acknowledge the help and support of my colleagues at the Centre for Paediatric Epidemiology and Biostatistics, Institute of Child Health, particularly Lucy Pembrey, Kirsty England, Kirsty Little and Icina Shakes. I would also like to thank Rodolphe Thiébaut for his advice on implementing the left-censored models in SAS.

Finally, I take pleasure in thanking my wonderful family and friends for their support and enthusiasm, and Zee for her support and understanding and for being my pillar of strength.

I dedicate this work to the memory of my mother.

The European Collaborative Study is a concerted action of the European Commission (PENTA/ECS 018865). The Medical Research Council (UK) Sexual Health and HIV Research Strategy Committee provided support to the ECS coordinating centre.

Chapter 1 HIV infection in pregnancy

1.1 Introduction

By December 2006, more than 20 years into the human immunodeficiency virus-type 1 (HIV) epidemic, an estimated 39.5 million people worldwide was living with HIV (UNAIDS 2006). Forty-five percent of these cases are made up of women; a demographic group in which the incidence of HIV infection is increasing (UNAIDS 2006). There were an estimated 4.3 million adults and children newly infected with HIV in 2006, almost half a million more than in 2004, with the substantial majority of these new infections occurring in sub-Saharan Africa (UNAIDS 2006). A small subset, 24 000, of these new infections relates to adults and children living in Western and Central Europe.

The rate of newly diagnosed cases of HIV infections reported in 2005 varied widely between the three sub-regions of Europe: the highest rate was observed in the East (186 per million population), over twice that reported in the West (82 per million) and twenty times that in Central Europe (9.4 per million) (EuroHIV 2006). Similarly, the rates of new infections and prevalence of HIV varies widely between countries (EuroHIV 2006;Hamers and Downs 2004), with estimates of newly diagnosed infections reported by the end of 2005 ranging from as low as 16.9 per million in Poland to as high as 242.5 per million in Ukraine (EuroHIV 2006).

1.2 Epidemiology of HIV infection

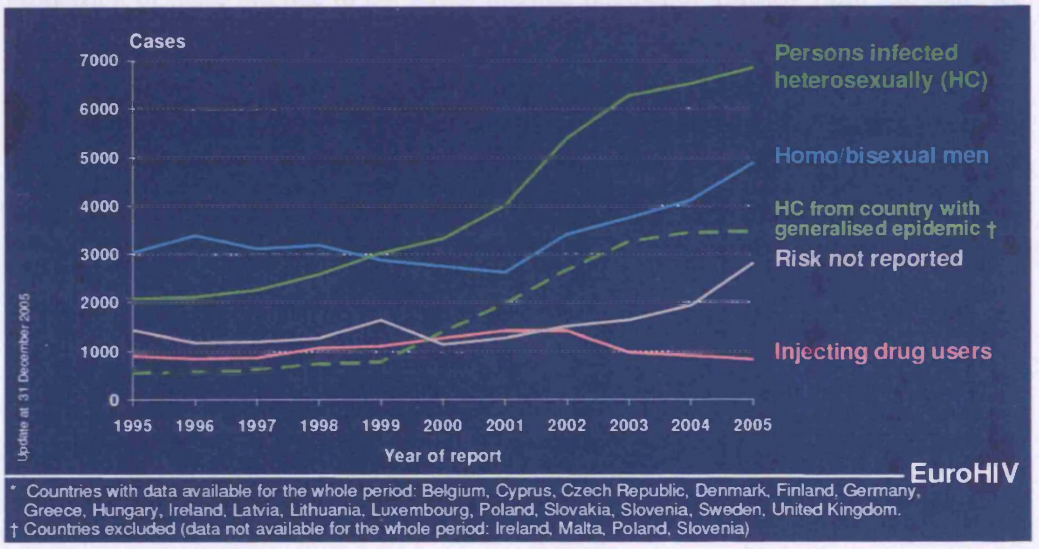
By the time the first cases of acquired immune deficiency syndrome (AIDS) were reported in the early 1980s (Centers for Disease Control 1981;Barre-Sinoussi *et al.*

1983), HIV had already started to spread among homosexual and drug-user communities throughout Western Europe. It has been estimated that HIV incidence among homosexual men peaked in the period 1982-1984 and in 1987-1988 among injecting-drug users (Downs *et al.* 2000). The number of heterosexually acquired infections increased gradually during the late 1980s and early 1990s.

The epidemic has changed in more recent times and since the late 1990s there has been a gradual decrease in HIV diagnoses among injecting drug users and a substantial increase in people infected through heterosexual contact in Europe (Figure 1.1). Increasing heterosexual transmission in Western Europe has seen the proportion of new HIV diagnoses among women increase from 25% in 1997 to 38% in 2002; in 2005 heterosexually-acquired infections accounted for 56% of new HIV diagnoses, with over half of these cases being among women (Hamers and Downs 2004;EuroHIV 2006). Figure 1.1 also shows the rising number of cases relating to heterosexual contact with people originating from countries with a generalised HIV epidemic, and in 2005 an estimated 60% of heterosexually-acquired infections were attributed to cases from a country with a serious HIV epidemic, particularly sub-Saharan Africa (EuroHIV 2006).

Although new HIV diagnoses in Central Europe are through similar routes to the West (albeit rarely through heterosexual contact with persons originating from countries with serious HIV epidemics), the story in Eastern Europe is very different; nearly 90% of all HIV cases in 2005 were reported from only two countries, the Ukraine and the Russian Federation, with injecting drug use the predominant transmission group (63%) in this region (EuroHIV 2006).

Figure 1.1 HIV infections newly diagnosed by transmission group 1995-2005, European Union.



Source: EuroHIV 2006, http://www.eurohiv.org/.

The prevalence of HIV infection in pregnant women varies across Europe with estimates of antenatal prevalence ranging from a high of 0.48% in Estonia (2002 data) and >0.6% in Ukraine to as low as <0.03% among women in Finland (EuroHIV 2005). Intermediate levels (0.03-0.3%) of HIV prevalence were reported among women giving birth in Germany, Italy, Spain and the United Kingdom and among pregnant women in the Netherlands and Norway (EuroHIV 2005). A distinguishing trend in Europe has been the continuing increase since 2000 in prevalence among women giving birth: the prevalence of HIV among pregnant women living in England has increased from 0.11% in 2001 to 0.19% in 2004 and 0.22% in 2005, reaching levels similar to those found in eight regions of Spain (EuroHIV 2005;Health Protection Agency 2006). Rapid increases of HIV prevalence among pregnant women have been reported in Estonia (a five-fold increase since 2000) and Ukraine (doubling of prevalence since 2000, reaching 0.34% in 2004) (EuroHIV 2005). Estimates of antenatal prevalence also vary within countries, for example the prevalence in London (0.44%) is almost four times greater than in the

rest of the country (0.13%) (EuroHIV 2005;Health Protection Agency 2006). A large proportion of these are due to the larger migrant population (mostly from sub-Saharan Africa); in 2005 antenatal prevalence of HIV stood at 2.4% among women born in sub-Saharan Africa, compared to 0.04% among UK-born women (Health Protection Agency 2006).

The rates published by EuroHIV and the Health Protection Agency are based on surveillance data, a key instrument in monitoring the HIV epidemic in Europe and the UK. However, a major limitation is that they are based on reported HIV diagnoses and unlikely to be representative of HIV incidence (EuroHIV 2004;Hamers and Downs 2004). The patterns of reporting HIV diagnoses may have changed over time and may include infections that occurred several years previously and depend on the uptake of HIV testing. Additionally, comparisons made between countries based on these reported rates assume that the quality and coverage of national surveillance are comparable which may not always be the case. These factors should be taken into account when interpreting surveillance figures (EuroHIV 2004;Hamers and Downs 2004).

The European Collaborative Study (ECS), covering 23 regional centres in 10 countries, recently carried out an epidemiological analysis on 6,000 mother-child pairs to compare the HIV mother-to-child transmission epidemic between Western and Eastern Europe and found that in Western European centres between 1985-1989, the majority of HIV-infected women were white (93%) and mainly injecting drug users (79%) with a third aware of their infection before the time of conception (European Collaborative Study 2006). The ECS results reflect the trends described above in the general HIV population and in the period 2000-2004 the number of black women had increased to 42% (mostly

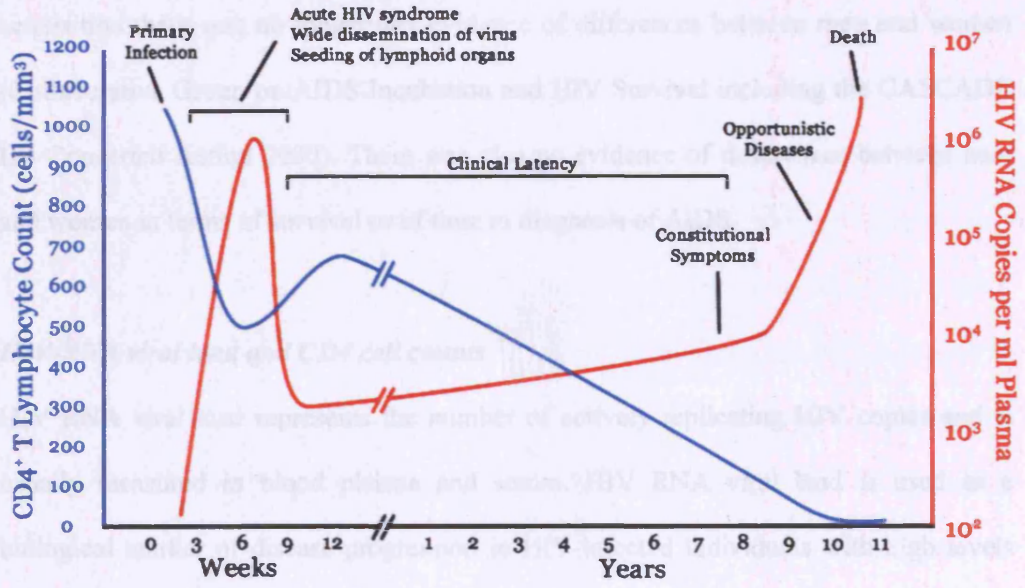
originating from sub-Saharan Africa) and the number of infections acquired through injecting drug use (IDU) had decreased to 20% with a concurrent increase in heterosexually-acquired infections (European Collaborative Study 2006). Additionally, an increase in the median maternal age from 21.7 years to 31.7 years between these two periods was observed. HIV-infected pregnant women enrolling in the Ukraine ECS centres in the period 2000-2004 were more similar to women enrolled in the earlier period in Western Europe; mostly white (98%), of young maternal age (median 25.4 years) and a large proportion acquiring HIV through injecting drug use or heterosexual contact with injecting drug users.

1.3 HIV infection and disease progression in adults

HIV is a Ribonucleic acid (RNA) virus, which targets the cell-associated immune system, particularly CD4 cells (Ho *et al.* 1995). Following primary infection, there is a period characterised by a lack of clinical symptoms, but accompanied by increasing virus and depleting levels of CD4 cells (Figure 1.2).

The main (M) group of HIV is responsible for most HIV infections in the global pandemic and is subdivided into nine subtypes (or clades) (Thomson *et al.* 2002). The most prevalent are subtypes B (found mainly in North America and Europe), A and D (found mainly in Africa), and C (found mainly in Africa and Asia). Increased HIV diversity has implications for diagnostic and treatment efficacy although data remain limited and most current HIV research remains focused on subtype B (Perrin *et al.* 2003).

Figure 1.2 A generalized graph of disease progression and the relationship between HIV RNA copies and CD4 cell counts over the average course of untreated HIV infection



Source: (Luzzi *et al.* 2003)

Without antiretroviral therapy (ART), immunosupression and the onset of AIDS are inevitable and the risk of opportunistic infection and other serious clinical symptoms becomes great. AIDS is diagnosed according to specific clinical and immunological criteria (Centers for Disease Control 1992) and left untreated the median time from seroconversion to the onset of AIDS is estimated to be 11 years (Collaborative Group on AIDS Incubation and HIV Survival including the CASCADE EU Concerted Action 2000) and from AIDS to death around 2 years (Mocroft *et al.* 1997). A meta-analysis of 38 studies of HIV-infected individuals before widespread use of ART found time since seroconversion and age at seroconversion to be major determinants of survival and development of AIDS in Europe, although AIDS incidence varied significantly by mode of acquisition (Collaborative Group on AIDS Incubation and HIV Survival including the CASCADE EU Concerted Action 2000).

A female-to-male mortality rate ratio of 0.77 was also reported, but after restricting the analysis to the two main exposure categories of injecting drug use and heterosexual acquisition there was no longer any evidence of differences between men and women (Collaborative Group on AIDS Incubation and HIV Survival including the CASCADE EU Concerted Action 2000). There was also no evidence of differences between men and women in terms of survival or of time to diagnosis of AIDS.

HIV-RNA viral load and CD4 cell counts

HIV RNA viral load represents the number of actively replicating HIV copies and is usually measured in blood plasma and serum. HIV RNA viral load is used as a biological marker of disease progression in HIV-infected individuals with high levels indicating primary infection or advanced disease progression (Figure 1.2) and is an important predictor for the risk of AIDS (Mellors *et al.* 1997;Sabin *et al.* 2000;Sterling *et al.* 2001). CD4 T-lymphocytes are the main target of HIV and hence their number declines as HIV infection progresses (Figure 1.2); they are therefore often used to monitor the extent of immune suppression in HIV-infected individuals and as an indicator of when to start or modify treatment.

The number of HIV RNA copies is assessed with the use of a variety of laboratory assays, which have a lower limit below which quantification is not possible. Assays today generally have a lower limit of detection between 500 and 20 copies/ml, while the older, less sensitive versions, had higher thresholds for detection. at between 2000-4000 copies/ml. Viral load measurements below the specific threshold value for each assay are therefore left-censored at the detection limit, and the distributions of HIV RNA viral load over time tend to be positively skew.

A large retrospective analysis of men with HIV carried out by the Multicenter AIDS Cohort Study (MACS) found higher levels of HIV RNA to be associated with larger decreases in CD4 count per year, with HIV RNA levels >30,000 copies/ml associated with a mean decrease in CD4 count of 70.5- 82.9 cells/mm^3 compared to a mean decrease of 30.4-42.3 cells/mm^3 among men with HIV RNA levels $\leq$500 copies/ml (Mellors *et al.* 1997). A study of 107 mostly untreated infected men observed for up to 17 years in the UK found little evidence of a long-term period over which HIV RNA levels remain stable and in most cases levels increased soon after seroconversion; mean HIV RNA levels were estimated to be 3.26 $\log_{10}$ copies/ml 1 year after seroconversion and estimated to increase by 0.11 $\log_{10}$ copies/ml each year thereafter (Sabin *et al.* 2000).

Studies on untreated men and women have shown lower HIV RNA viral loads to be associated with black ethnicity (Anastos *et al.* 2000;Saul *et al.* 2001) and the authors of these studies postulate ethnic differences in HIV clades, HLA haplotypes and other social and biologic factors as possible explanations for these findings. Some studies have also shown an association between persons with a history of IDU and lower viral load (Anastos *et al.* 2000;Touloumi *et al.* 2004), although an analysis from the Woman and Infant Transmission Study found no association between hard drug use (including IDU) and HIV RNA levels (Thorpe *et al.* 2004).

The aforementioned MACS (Mellors *et al.* 1997) compared the prognostic value of HIV RNA levels and CD4 lymphocytes and showed that the former marker was a better prognostic indicator of HIV disease progression than CD4 cell count. However, it concluded that a more accurate prognosis was only possible through the combination of

the two markers. More recent studies have confirmed this through the use of bivariate models of these two markers, which have revealed better fits to the data than two separate univariate models (Boscardin *et al.* 1998;Thiébaut *et al.* 2002;Thiébaut *et al.* 2005).

1.4 Treatment of HIV infection with antiretroviral therapy (ART)

Zidovudine (ZDV), a nucleoside reverse-transcriptase inhibitor (NRTI) licensed in 1987, was the first antiretroviral (ARV) drug used for the treatment of HIV infection (Fischl *et al.* 1987). Additional NRTIs and two further drug classes, non-nucleoside reverse transcriptase inhibitors (NNRTI) and protease inhibitors (PI) were subsequently developed and the principal drug names for these classes are given in Table 1.1 (AIDSinfo 2007).

Table 1.1 Principal antiretroviral agents by drug class

Nucleoside reverse transcriptase inhibitors (NRTI)	Non-nucleoside reverse transcriptase inhibitors (NNRTI)	Protease inhibitors (PI)
Zidovudine (AZT)	Efavirenz (EFV)	Ritonavir (RTV)
Didanosine (ddI)	Nevirapine (NVP)	Indinavir (IDV)
Lamivudine (3TC)	Delavirdine (DLV)	Nelfinavir (NFV)
Stavudine (d4T)		Saquinavir (SQV)
Abacavir (ABC)		Atazanavir (ATZ)
Zalcitabine (DDC)		Amprenavir (APV)
Tenofovir (TDF)		Fosamprenavir (FSV)
		Lopinavir (LPV)[†]

[†] - LPV is co-formulated with low-dose RTV

These different drug classes target the viral replication cycle at different stages to disturb replication of the virus, e.g. PIs disturb the cycle by inhibiting the activity of the

enzyme, protease. More recently, enfurtivide (T-20), a fusion inhibitor (FI), has been developed and additional new therapies for HIV-1 infection, such as the CCR5 antagonists, are now in advanced stages of clinical development (Repik *et al.* 2007).

Subsequent evidence indicated that the benefits of ZDV monotherapy in HIV-infected individuals were small, and short-lived (Fischl *et al.* 1989). Consequently, two large randomised controlled trials carried out in 1996 demonstrated that initiation of treatment with a combination of two ARVs (dual therapy) delayed disease progression for longer than treatment with ZDV alone (Delta Study 1996;Hammer *et al.* 1996). Highly active antiretroviral therapy (HAART) generally consists of at least three potent antiretroviral drugs, usually involving at least one PI or NNRTI in addition to a NRTI 'backbone'. HAART was introduced in 1996 and was found to delay disease progression, reduce HIV replication and increase CD4 cell counts further than with two drug combinations only (Collier *et al.* 1996). The use of HAART in developed countries is now widespread and is the standard of care for the HIV-infected population in this setting (Gazzard *et al.* 2006;DHHS Panel on Antiretroviral Guidelines for Adults and Adolescents 2006). Additionally, guidelines now advise that PI-based HAART regimens (usually including the PIs SQV or FSV) be 'boosted' with the PI RTV, as they have been shown to be superior to PIs alone (Walmsley *et al.* 2002).

1.5 Mother-to-child transmission of HIV

The first cases of AIDS in children were reported in 1982 (Centers for Disease Control 1982); in 2003 alone, 630,000 children acquired HIV, mostly through mother-to-child transmission (MTCT). MTCT can occur during pregnancy, at the time of delivery or postnatally through breastfeeding (Newell *et al.* 2002). Intra-uterine transmission may

occur as a result of fetal exposure to cell-free and/or cell-associated HIV in amniotic fluid and through infection of the placental cells and/or disruption of the integrity of the placenta (as evidenced by the association of an increased risk of MTCT with chorioamnionitis (Mofenson *et al.* 1999) and amniocentesis (Mandelbrot *et al.* 1996)). Mechanisms of intrapartum transmission include microtransfusions between maternal and fetal blood circulation and the ascension of HIV infection from the vagina and cervix to amniotic fluid and the fetus after rupture of membranes (Mwanyumba *et al.* 2002;Tuomala *et al.* 2003). A substantial amount of intrapartum transmission may occur through direct contact between the infant passing through the birth canal and infective cervicovaginal secretions and blood. Prior to therapeutic interventions, reported MTCT rates ranged from 15 to 20% in Europe, 13-25% in the USA, 25-40% in Africa to 13-48% in South and South East Asia (European Collaborative Study 2001;The Working Group on Mother-To-Child Transmission of HIV 1995). These geographical differences can be attributed to varying infant feeding practices and underlying population characteristics.

1.5.1 Prevention of mother-to-child-transmission (PMTCT)

In Europe today, effective interventions used in the prevention of MTCT (PMTCT) include elective caesarean section (CS), antenatal ART, perinatal and neonatal prophylaxis and the avoidance of breastfeeding. Table 1.2 highlights the trends in strategies for PMTCT in Europe and the associated reductions in MTCT rates over time.

Maternal plasma HIV RNA level is the strongest individual risk factor for MTCT (European Collaborative Study 1999;Cooper *et al.* 2002;Ioannidis *et al.* 2001;Mofenson *et al.* 1999). The ECS showed that the risk of MTCT was doubled for every unit

increase in $\log_{10}$ RNA with a predicted MTCT rate of 45% at 1,000,000 copies/ml, compared with 8.7% at 1000 copies/ml (European Collaborative Study 1999). It has also been shown that primary infection during pregnancy, severe immunodeficiency and symptomatic HIV infection, themselves associated with high viral loads, increase the risk of transmission (Ioannidis *et al.* 2001;Cooper *et al.* 2002;Mofenson *et al.* 1999). For these reasons, preventative measures to reduce MTCT generally focus on, but are not limited to, reducing HIV RNA levels by the time of delivery using ART and HAART. Obstetric factors which increase the risk of MTCT include vaginal delivery and prolonged duration of rupture of membranes (European Collaborative Study 1999;International Perinatal HIV Group 2001;European Collaborative Study 2005b). The use of elective CS delivery has been associated with a two-thirds reduction in the risk of transmission, compared to that with vaginal deliveries (European Collaborative Study 1999;European Mode of Delivery Collaboration 1999). In an analysis of the ECS, emergency CS delivery was associated with a 25% reduction in risk, however, there were no cases of transmissions in children who were delivered by emergency CS with intact membranes (European Collaborative Study 2005b); the International Perinatal HIV Group estimated a 2% increase in the risk of MTCT with every hour increase of duration of ruptured membranes in women delivering by vaginal or emergency CS delivery (International Perinatal HIV Group 2001).

Table 1.2 Prevention of MTCT in Europe

1985	• **First report indicating possible MTCT through breast milk (Ziegler *et al.* 1985)**
Pre-1994	• **No antiretroviral prophylaxis for the prevention of MTCT** • Limited data in the early 1990s indicate a 14% additional risk of transmission from breast milk, over and above that of risk in utero and during delivery (Dunn *et al.* 1992) • Avoidance of breastfeeding recommended (World Health Organisation 2005) • MTCT rates in Europe 15-18% (ECS 1996;Gabiano *et al.* 1992)
1994	• **Results of the ACTG 076 randomised placebo-controlled trial (Connor *et al.* 1994)** • Prophylactic three-part regimen of ZDV monotherapy • Antenatal oral ZDV started between 14 and 26 weeks' gestation • Intravenous ZDV during labour • Neonatal ZDV for 6 weeks postpartum • MTCT rate at 18 months: 8.3% in the ZDV group and 25.5% in the placebo
1994-1996	• **Clinical practice quickly changed in developed countries** • Rapid uptake and widespread use of ZDV monotherapy for prevention of MTCT (ECS 1998;Cooper *et al.* 1996) • MTCT rate in Europe of 5-8 % for non-breastfeeding population (European Collaborative Study 1998;Mayaux *et al.* 1997b)
1997	• **More potent antiretroviral regimens (including HAART) become available and are increasingly prescribed among adults in Western Europe** • Limited use of HAART in pregnancy at this time
1999	• **Results of the European Mode of Delivery Collaboration randomised trial (European Mode of Delivery Collaboration 1999)** • Women randomised to elective CS or vaginal delivery • Elective CS delivery before labour and before rupture of membranes associated with a two-thirds reduction in the risk of MTCT
2002	• **Results of the PACTG 316 placebo-controlled trial (Dorenbaum *et al.* 2002)** • Included women receiving any antenatal ART regimens excluding NNRTI class of drugs (mono, dual or HAART) • Additionally, oral single-dose nevirapine to women after onset of labour and postpartum to neonate • MTCT rate of 1.4% in nevirapine group and 1.6% in placebo group • No additional benefit from intrapartum/neonatal nevirapine • However, confirmed effectiveness of antenatal ART • **European consensus guidelines on pregnancy and HIV infection updated (Newell *et al.* 2002)** • ZDV-containing HAART, with type and timing of regimen dependent on clinical status • Intrapartum intravenous ZDV and neonatal ZDV for 4-6 weeks • Women recommend to have an elective CS • Avoidance of breastfeeding
2002-2005	• **Antenatal HAART use widespread among Europe (Italian Register for Human Immunodeficiency Virus Infection in Children 2002;European Collaborative Study 2005b):** • 85% of women enrolling in the ECS in 2003 receiving HAART pre-pregnancy (ECS 2005b) • MTCT rates of less than 2% reported in Europe (ECS 2005b)

On the other hand, CS is also associated with a higher rate of post-partum complications among HIV-infected women than vaginal deliveries (Read *et al.* 2001a) and it has been suggested that elective CS to reduce the risk of MTCT may not carry an additional benefit for women receiving HAART with undetectable viral load near the time of delivery (Boer *et al.* 2007). Nevertheless, findings from several observational studies have confirmed the independent effect of elective CS in reducing the risk of MTCT for women receiving HAART with low viral loads (Ioannidis *et al.* 2001;Cooper *et al.* 2002;European Collaborative Study 2005b). European guidelines recommend that elective CS should be scheduled for 38 weeks gestation, rather than 39 weeks, to avoid the initiation of labour or the rupture of membranes (Newell *et al.* 2002;British HIV Association 2005).

1.5.2 ART in PMTCT

ART reduces the risk of MTCT by decreasing viral replication, and thus plasma viral load in the woman. Additionally, for drugs which cross the placenta, mostly the reverse transcriptase inhibitors such as ZDV, nevirapine (NVP) and lamivudine (Mirochnick 2000;Mandelbrot *et al.* 2001b), they act as post-exposure prophylaxis of the neonate during and after exposure to the virus. The results of the AIDS Clinical Trial Group (ACTG) 076 trial in 1994 showed the efficacy of ZDV use in PMTCT and marked the beginning of the widespread use of ART during pregnancy (Table 1.2). There have since been many advances in the therapeutic management of HIV disease and MTCT in Europe (Table 1.2), notably the introduction of HAART.

Although, it has not been possible to evaluate the effectiveness of antenatal HAART in the prevention of MTCT in randomised placebo-controlled trials, its effectiveness has

been confirmed with data from observational studies (Cooper *et al.* 2002;European Collaborative Study 2005b). In the Women and Infants Transmission Study, the MTCT rate was 20.0% for women not receiving any antenatal ART, 10.4% for women receiving ZDV monotherapy, 3.8% for women receiving dual-therapy and 1.2% for women on HAART (Cooper *et al.* 2002). The Italian Register also reported a decline in the transmission rate over time, from 15.5% in the period 1985-1995 to 2.4% in the period 1996-1999, when combination prophylactic ART and elective caesarean delivery were combined (Italian Register for Human Immunodeficiency Virus Infection in Children 2002). The ECS has reported similar trends, and with the appropriate prophylactic interventions in place, MTCT rates are now as low as 1% (European Collaborative Study 2005b;European Collaborative Study 2006).

1.6 HIV in pregnancy

Pregnancy is associated with relative immunosuppression to protect the fetus from maternal rejection (Wegmann *et al.* 1993;Castilla *et al.* 1989). This pregnancy-associated decline in immune responses has also been associated with an increased frequency of infectious diseases, including viral infections such as influenza and varicella (Rich *et al.* 1995;Pickard 1968;Greenberg *et al.* 1958) and with the modification of the course of diseases with an immunological component, such as systemic lupus erythematosus (Tincani *et al.* 1991). This has raised concern that pregnancy itself might also impact on disease progression in HIV-1-infected women.

A systematic review and meta-analysis of studies of ART-naïve women between 1983 and 1996 found a significant association between the effect of pregnancy and disease progression, albeit not a strong one and a further sensitivity analysis revealed this

association to be significantly less common in developed countries (French and Brocklehurst 1998). More recent studies carried out have found no deleterious effect of pregnancy on HIV disease progression (Weisser *et al.* 1998;Bessinger *et al.* 1998;Saada *et al.* 2000), but focused mainly on clinical outcomes. Given that CD4 cell counts, are extremely informative of immune depression, several studies have been carried out to examine the impact of pregnancy on levels of CD4 counts in HIV-infected women (ECS and Swiss Cohort and European Collaborative Study and the Swiss HIV Pregnancy Cohort 1997;Ibrahim *et al.* 2004;Burns *et al.* 1996)

1.6.1 CD4 cell counts during pregnancy

A joint analysis by the ECS and the Swiss HIV Pregnancy Cohort investigated immunological markers of disease progression in more than 400 untreated pregnant women (ECS and Swiss Cohort and European Collaborative Study and the Swiss HIV Pregnancy Cohort 1997). The pattern of CD4 counts was shown to change non-linearly with time during and after pregnancy; the mean CD4 count was seen to decline to its lowest level six months prior to delivery and subsequently increased until delivery; a sharp increase in the average CD4 count was subsequently observed until three months post-delivery, with the average dropping to baseline six months postpartum. CD4 percentage, expressed as a percentage of absolute lymphocyte counts, was stable during pregnancy, suggesting that pregnancy-related changes in total lymphocyte counts may be responsible for the changes seen in CD4 counts. (ECS and Swiss Cohort and European Collaborative Study and the Swiss HIV Pregnancy Cohort 1997). Additionally, a South African cross-sectional study comparing lymphocyte changes in pregnancy and postpartum between HIV-infected and uninfected women found that that CD4 counts did not decrease significantly during pregnancy; however, among the

infected women, there was a significant increase in postpartum CD4 cell counts compared to those in the third trimester (Ibrahim *et al.* 2004). It has been suggested that haemodilution, an increase in the volume of cell-free solution, during a normal pregnancy leads to a decrease in the number of CD4 lymphocytes (Castilla *et al.* 1989); the mechanisms behind changes in the cellular immune system in pregnancy are not well understood and the relevance for women with HIV-infection is unclear.

1.6.2 HIV-RNA viral load during pregnancy

As the most important predictor for MTCT and disease progression, a number of studies have examined HIV RNA levels during pregnancy and postpartum and these have been summarised in Table 1.3. Most of these studies revealed stable levels of HIV RNA during pregnancy followed by varying degrees of increases in HIV RNA levels postpartum. The stable levels during pregnancy would suggest that any deleterious effect of a decline in certain immune responses during pregnancy is offset by other `immunologic changes, e.g. alterations in cytokines, which are important immunomodulators that reportedly play an important role in modifying HIV replication (Fauci 1996). Physiologic and hormonal changes during pregnancy, such as elevated oestrogen/progesterone levels and haemodilution, could also impact on the stimulation of HIV replication; such elevated hormone levels are reversed after delivery leading to increased levels postpartum. It has also been suggested that the observed increases in viral load postpartum in treated women could be due to the emergence of viral resistance and poorer adherence (Watts *et al.* 2003;Ickovics *et al.* 2002).

Table 1.3 Features of studies investigating HIV RNA levels in pregnancy and postnatally

Author (yr)	Time period	Data source	#s with viral load	Statistical Methods	Treated/ Prophylaxis	Summary of results	Limitations
Mayaux et al. (97)	89-94	SEROGEST, France	320	Paired Wilcoxon rank-sum	21% ZDV (removed from analysis after initiation)	-No significant variation during pregnancy, postpartum values lower -Highest HIV RNA level observed during pregnancy lower in African women	- Did not adjust for confounders - Ignored repeated measures
Cao et al. (97)	93-97	ARIEL, USA	204	Paired Wilcoxon rank-sum and linear regression	85% ZDV	- -2 copies/ml a day decrease in HIV RNA during pregnancy compared to increase of 6 copies postpartum - did not vary during pregnancy in regression	- Mostly treated women -Ignored repeated measures -only measures from third trim and delivery
Melvin et al. (97)	91-95	NW Family center, Washington, US	44	Paired Wilcoxon rank-sum and Mann-Whitney U	61% ZDV	-Stable HIV RNA during pregnancy and at 6 weeks postpartum in both ZDV and untreated groups	- Small numbers - Did not adjust for confounders -Ignored repeated measures
O'Shea et al. (98)		Europe	38	Paired comparisons	12.6% ZDV	-For 19 untreated women, relatively stable HIV RNA during pregnancy, no consistent pattern in direction of change	- Small numbers - Did not adjust for confounders -Ignored repeated measures

Table 1.3 contd.

Author (yr)	Time period	Data source	#s with VL	Statistical Methods	Treated/ Prophylaxis	Summary of results	Limitations
Burns et al. (98)	86-95	Mothers and Infants Cohort Study, USA	160	Longitudinal random effects model	4% ZDV during pregnancy, 26% ZDV or DDI postpartum	- Overall HIV RNA rose significantly during study period (early pregnancy to 2 years postpartum) - however, slope of HIV RNA during pregnancy and 1st year postpartum not different from zero	-only 213 measurements during pregnancy
Watts et al. (03)	93-97	PACTG 185, USA	493	Repeated measures linear regression	86%	- Consistent but non-significant decrease in HIV RNA during pregnancy, followed by modest significant increase during 12 weeks postpartum	-mostly treated women
Ekpini et al. (02)	96-98	Abidjan, Côte d'Ivoire	49	Paired Wilcoxon rank-sum and Mann-Whitney U	69% ZDV between 3	-in untreated women, HIV RNA varied between -0.08 to 0.23 $\log_{10}$ copies/ml during pregnancy but not significantly different from baseline	-small numbers of untreated women - Did not adjust for confounders -Ignored repeated measures

However, the increases in viral load postpartum observed for both untreated and treated women (Cao *et al.* 1997;Watts *et al.* 2003) suggest that these are insufficient explanations for these increases.

Most studies of HIV RNA viral load and CD4 cell changes in pregnancy to date have included only small numbers of women and larger studies of both treated and untreated women are necessary to better understand the variation of these markers during pregnancy, particularly with regard to changes under HAART treatment. Importantly, studies of viral load during pregnancy did not report adjustment for maternal immune status, and as revealed in studies of non-pregnant adults (Mellors *et al.* 1997;Boscardin *et al.* 1998;Thiébaut *et al.* 2002;Thiébaut *et al.* 2005), the combination of HIV RNA viral load and CD4 cell counts is essential in understanding not only the true underlying relationship between the two markers, but also their association with other factors such as ART. However, to date there have been no studies to examine CD4 cell counts and HIV RNA viral load during pregnancy using appropriate bivariate statistical methods.

1.7 ARVs and adverse events in pregnant women

With increasing numbers of women becoming pregnant while receiving HAART or initiating treatment in the early stages of pregnancy (European Collaborative Study 2001;Cooper *et al.* 2002;Watts *et al.* 2004a), concerns have been raised relating to the potential impact on pregnancy outcome and safety issues for the mother and the exposed infant.

1.7.1 Pregnancy and neonatal outcomes

In 1998, a Swiss study assessing the safety of ART in pregnancy reported a 33% prematurity (<37 weeks) rate among newborns exposed to antenatal HAART (Lorenzi *et*

al. 1998). The association between antenatal HAART use and premature delivery has since been confirmed with data from several European observational studies (ECS and Swiss Cohort and European Collaborative Study and the Swiss HIV Pregnancy Cohort 2000;European Collaborative Study 2003;Townsend *et al.* 2007b). These findings are consistent with the hypothesis that an immunological mechanism may underlie the association between antenatal HAART and pregnancy duration, that is that prematurity is caused by the early onset of labour/delivery rather than being of fetal or uteroplacental origin. (Fiore *et al.* 2006). Although some US Studies have reported similar findings of an increased risk of prematurity with antenatal HAART use (Cotter *et al.* 2006;Schulte *et al.* 2007), others have not (Cooper *et al.* 2002;Tuomala *et al.* 2002;Tuomala *et al.* 2005). The discrepancy may be partly explained by differences in study populations (in terms of illicit-drug use as well as HIV infection) and by the lack of detailed data on prenatal ART use in the American studies (Thorne *et al.* 2003). A recent meta-analysis of published studies on HAART and prematurity in the US and Europe reported a significant degree of heterogeneity between studies confirming these differences (Kourtis *et al.* 2007); however further work is needed to ascertain the source of these differences.

In an analysis of the ECS, involving 319 infants with in utero HAART exposure out of a total of 2326 mother-child pairs, HAART without or with a PI was associated with a 2.66 and 4.14-times greater risk of prematurity respectively, compared to those with no exposure, after adjusting for maternal CD4 count, age and IDU (European Collaborative Study 2003). Recent studies have reported conflicting findings on whether antenatal PI-based HAART regimens are associated with an increased risk of prematurity compared to HAART without a PI (Cotter *et al.* 2006;Schulte *et al.* 2007).

In light of the finding of an association between HAART use during pregnancy and an increased risk of prematurity, some European centres have operated a policy of scheduling elective CS deliveries at 36-37 weeks to avoid initiation of labour or rupture of membranes and emergency CS delivery (Newell *et al.* 2002).

The use of HAART during pregnancy, particularly in the early stages when organogenesis is taking place, has raised concerns relating to teratogenicity. Although human data have shown no evidence of an increased risk of congenital abnormalities associated with ZDV prophylaxis (European Collaborative Study 2003), most of the ARV drugs in these regimens have not been tested extensively in humans for teratogenic potential. However, efavirenz has been associated with teratogenicity in animal studies (Nightingale 1998) and there have been reports of neural tube defects in children with neonatal exposure to efavirenz in the first trimester of pregnancy (De Santis *et al.* 2002). It has therefore been recommended that efavirenz be avoided in the first trimester of pregnancy and that women of childbearing age be offered pregnancy testing before initiating therapy with this drug, and counselling on the need to be on effective contraception to avoid pregnancy throughout treatment (Public Health Service Taskforce 2006b).

Although limited to a small sample size including only few women initiating treatment in early pregnancy, a retrospective study found an increased risk of malformations in infants exposed to *PneumoCystis jiroveci* pneumonia (PCP) prophylaxis and combination of ART in the first trimester (Jungmann *et al.* 2001). A European study focusing on the health of uninfected infants exposed to ART in-utero, observed a 1.4% prevalence of congenital abnormalities in those infants ARV-exposed (mainly to mono/dual therapy), similar to the 1.6% seen in those not exposed (European

Collaborative Study 2003). The National Study of HIV in Pregnancy and Childhood, a paediatric and obstetric surveillance system in the United Kingdom and Ireland, reported relatively higher rates of congenital abnormalities than the ECS, but did not find a significant difference between children exposed to ART in-utero (3.4%) and those who were not (2.2%). There was also no increased risk associated with exposure in the first trimester, compared to exposures in the second or third trimester of pregnancy (Townsend *et al.* 2006).

The Antiretroviral Pregnancy Registry (APR), based in the United States, was established by the pharmaceutical industry in 1988 in order to monitor exposures to ART during pregnancy and to assess the risk of major birth defects (Antiretroviral Pregnancy Registry Steering Committee 2004). The APR has not identified any specific teratogenic risk associated with any particular class of ARV drugs with data on 3583 live births reported to date (Antiretroviral Pregnancy Registry Steering Committee 2004;Covington *et al.* 2004;Watts *et al.* 2004b). In addition, there was no significant difference in the risk of congenital abnormalities between first trimester exposure to ART and later exposures (Watts *et al.* 2004b). However, data from case report registries such as the APR have limitations, such as sample size restrictions, potential classification and reporting bias, potential ascertainment bias and generalisability to the population of pregnant women with ART exposure worldwide; these factors need to be considered when interpreting the results of studies from the APR.

1.7.2 Risk for pregnant women

Results from some studies have also raised concerns about the safety of using certain classes of ARVs and specific drugs in pregnant woman with regard to their own health (Hitti *et al.* 2004;Sarner and Fakoya 2002;Mandelbrot *et al.* 2003;Food and drug

administration 2005;British HIV Association 2005;Coll O *et al.* 2005;Watts *et al.* 2004a).

Severe, and in some cases fatal, hepatotoxicity has been reported in non-pregnant HIV-infected patients receiving NVP-containing HAART (Baylor and Johann-Liang 2004). Symptoms range from hepatitis to hepatic failure and also include hypersensitivity skin reactions. Women are at a greater risk than men of developing such adverse reactions, which occur during the initial weeks of therapy, although the risk may continue for the duration of exposure (Mazhude *et al.* 2002;Bersoff-Matcha *et al.* 2001). There appears to be a strong association between the level of immunosuppression of the individual at the time of initiation of treatment and the likelihood of NVP -related hepatotoxicity (with women with CD4 cell counts >250 cells/mm^3 and men with >350 cells/mm^3 at increased risk than women and men with lower counts) (Stern *et al.* 2003). In the Pediatric AIDS Clinical Trials Group 1022, pregnant women were randomised to receive either NVP or nelfinavir, with lamivudine plus ZDV (Hitti *et al.* 2004), however the study was suspended after the enrolment of only 38 women due to greater than expected toxicity in the NVP group of 29% vs. 5% in the nelfinavir group; all the women with adverse events in the NVP arm had baseline CD4 counts >250 cells/mm^3 and none had a history of hepatitis B or C. However, there is conflicting evidence on the relationship between baseline CD4 counts and NVP-related hepatotoxicity (Manfredi *et al.* 2005;Manfredi and Calza 2006). Although data are lacking on whether or not the use of NVP-containing HAART during pregnancy is associated with an increase in the risk of hepatotoxicity, US and WHO recommendations express caution in the use of long-term NVP-containing HAART for pregnant women with CD4 cell counts 250-350 cells/mm^3 (Public Health Service Taskforce 2006a;World Health Organisation 2006).

There have also been case reports of lactic acidosis in pregnant women receiving HAART including a combination of stavudine and didanosine from before conception; these included some maternal deaths (Sarner and Fakoya 2002;Mandelbrot *et al.* 2003). It is therefore recommended that this drug combination is avoided during pregnancy (Food and drug administration 2005;British HIV Association 2005). Additionally, there have been reports of an increased risk of pre-eclampsia among women on HAART (Wimalasundera *et al.* 2002;Coll O *et al.* 2005) and also of glucose intolerance and gestational diabetes among women on HAART, particularly those receiving PIs (Watts *et al.* 2004a).

The risk of emergence of viral resistance to ARV drugs is increased when viral replication is only partially suppressed, for example, with sub-optimal therapies or with poor adherence to treatment regimens. The risk is also present for women who commence ART in pregnancy to reduce the risk of MTCT (and who do not require therapy for their own health) and discontinue therapy postpartum. Emergence of viral resistance is particularly common for those drugs where a single mutation is associated with resistance, such as lamivudine and NVP (Mandelbrot *et al.* 2001a;Lyons *et al.* 2005;Overton *et al.* 2005). The use of NVP -containing HAART in preventing MTCT has been associated with high levels of resistance (Lyons *et al.* 2005;Overton *et al.* 2005). A recent study in Ireland reported 7 primary mutations among five (13%) ART-naïve women who received therapy during pregnancy; all of these women received NVP -containing HAART (Lyons *et al.* 2005). Another study reported the finding of an association between the use of ZDV plus lamivudine for more than 8 weeks in pregnancy and a 50% risk of developing resistance (Mandelbrot *et al.* 2001a).

1.8 Statistical methods

Observational studies of HIV-infected pregnant women with prospective follow-up provide repeated measurements on HIV RNA viral load and CD4 cell counts which are non-independent due to within-subject correlation. Random effects are often used to deal with repeated measures (Laird and Ware 1982), but there are several additional issues in the analysis of longitudinal data of HIV markers. These include left-censoring of HIV RNA viral load, non-uniform intervals for measurements of CD4 cell counts and HIV RNA viral load, correlation between these HIV markers and determination of the functional form of non-linear patterns over pregnancy.

HIV RNA viral load data are subject to left-censoring where the true, unknown value is below the lower threshold level for detection. Moreover, with the widespread use of antenatal HAART during pregnancy, the proportion of infected women with undetectable levels of HIV RNA has increased (European Collaborative Study 2001;Read *et al.* 2001a;Cooper *et al.* 2002). Relatively crude methods used in dealing with left-censoring of HIV RNA measurements include taking either the quantification limit or half of this limit; this is more problematic with older versions of assay systems, which have higher thresholds for detection between 2000-4000 copies/ml. These methods can lead to biased parameter estimates with inflated standard errors, and more advanced methods to account for the presence of left-censoring are preferable (Hughes 1999).

Observational data reflect clinical practice, in that patients are not always enrolled at the same time and the frequency of and interval between clinic visits may differ between and within patients. When there is a difference in the frequency of follow-up between treatment groups or other variable of interest, bias can be introduced; the use of interval-

censored methods to account for non-uniform measurement intervals will improve the accuracy of treatment comparison estimates (Lindsey and Ryan 1998).

Univariate linear mixed effects models of HIV RNA and CD4 cell count data are extremely informative, but do not appropriately allow for the intrinsic correlation between the two markers; bivariate modelling of these has been shown to provide a better fit to such data and can also allow for the incorporation of the left-censoring of HIV RNA data (Thiébaut *et al.* 2005). Additionally, the changes in HIV RNA levels and CD4 cell counts over the course of infection are not necessarily linear (Mellors *et al.* 1997;Sabin *et al.* 2000), particularly among those receiving HAART (Fitzgerald *et al.* 2002;Thiébaut *et al.* 2005;Thiébaut *et al.* 2006) and the use of piecewise, polynomial and cubic spline models (Naumova *et al.* 2001;Lyles *et al.* 1999;Gurrin *et al.* 2005) may be essential in capturing the true underlying functional form of these markers over pregnancy.

1.9 Key points

- There has been an increase in the number of heterosexually-acquired HIV infections in Europe with women accounting for over half of these cases in 2005 and an increase in the number of infected women from sub-Saharan Africa.
- Available data on whether increasing gestation may impact on HIV RNA are limited, based on mostly treated women and suggest that viral load remains stable during pregnancy, increasing in the early postpartum period.
- In the absence of clinical trials, data from observational cohorts have confirmed the effectiveness of antenatal HAART in PMTCT, but data are lacking regarding the optimal approach to therapeutic management in pregnancy.
- The natural dynamics of markers of HIV RNA and CD4 cell counts during pregnancy are poorly understood.
- Studies of HIV RNA and CD4 levels in treated women are limited, focusing mainly around the time of delivery.
- There is conflicting evidence on whether the use of HAART with a PI is associated with an increased risk of premature delivery compared to that associated with NNRTI-based HAART.
- There has been no evidence to date to suggest an association between antenatal HAART and an increased risk of congenital abnormalities.

Chapter 2 Aims and Methods

2.1 Rationale for this work

There have been two major developments in the care of HIV-infected pregnant women in recent years. Firstly, the rate of MTCT has decreased to levels as low as 1% in developed countries where HAART and other interventions are available (European Collaborative Study 2006). Secondly, the widespread use of HAART has changed the clinical course of HIV/AIDS with significant improvements in AIDS-free survival and quality of life (Porter *et al.* 2003;Mocroft *et al.* 2003). These developments help explain the observed increase in likelihood of further live births in HIV-infected women in recent years (European Collaborative Study 2005a;Bryant *et al.* 2007).

In the beginning of the epidemic, the focus of studies relating to HIV infection in pregnancy concentrated on elucidating the rates of and risk factors for MTCT and the implications of pregnancy on HIV disease status in the short or longer term were somewhat overlooked. With the routine use of HAART for PMTCT now commonplace in resource-rich settings, concerns that ARVs may be teratogenic, especially when used in the first trimester, and that their use in pregnancy may increase the risk of prematurity have shifted the focus to the unborn child. However, the use of ART in pregnancy serves the best interest of both mother and child by reducing maternal viral load and thereby reducing the risk of MTCT (European Collaborative Study 2005b;Mofenson *et al.* 1999) and also slowing maternal disease progression (Mellors *et al.* 1997). The HIV epidemic is increasingly affecting women of reproductive age (UNAIDS 2006), with a concurrent increase in the numbers of infected women taking highly potent, complex combinations of drugs throughout pregnancy, often initiated before or in early

pregnancy. The impact of these drugs and their various combinations on pregnancy outcome, maternal and neonatal health and immunological/virological outcomes in pregnancy is unclear. Over the past decade there have been major changes in the characteristics of the HIV infected pregnant population in Europe, with a shift in mode of acquisition from IDU to heterosexually acquired infection (particularly in women from sub-Saharan Africa) (Hamers and Downs 2004;EuroHIV 2004), and the interaction of these changes on treatment outcome have not been fully explored.

Few data are available to indicate whether the risk of congenital abnormalities is increased by first trimester exposure to ARVs and/or by the use of HAART. Clinical trials in humans typically exclude pregnant women and many drugs have few or no data available on their safety in pregnancy, including teratogenicity (Public Health Service Taskforce 2006b). European observational studies have shown an increased risk of premature delivery associated with antenatal combination ART (Lorenzi *et al.* 1998;ECS and Swiss Cohort and European Collaborative Study and the Swiss HIV Pregnancy Cohort 2000;European Collaborative Study 2003). Recent studies from Brazil and the US have also examined the impact of PI-based HAART on increased risk of prematurity, but mostly with reference to none or mono/dual therapy only (Szyld *et al.* 2006;Schulte *et al.* 2007;Cotter *et al.* 2006). Data are lacking on whether the use of PI-based HAART increases the risk of prematurity over and above that associated with NNRTI-based HAART.

Since the beginning of the epidemic, the relationship of pregnancy to immune and physiological alterations has given rise to questions about its effect on HIV disease, both in the short term (during gestation itself) and the effect it may have on HIV disease

progression subsequent to delivery. Data are lacking on the dynamics of virological and immunological markers during pregnancy both in the absence of treatment and in women on stable HAART. Understanding the dynamics of these markers in pregnancy will facilitate the interpretation of changes in these markers among HIV-infected pregnant women receiving HAART at conception, and may be useful for examining treatment response in pregnancy.

A substantial minority of women in developed settings are diagnosed following antenatal testing and started on antiretroviral drugs for delaying their own disease progression and/or prevention of MTCT for the first time in pregnancy (European Collaborative Study 2005b). To date, there have been no clinical trials in this setting to address the questions of when HAART should be started in pregnancy and which regimens are more effective for optimal viral response in this group of ARV-naïve pregnant women.

2.2 Aims and Objectives

Aims

This thesis aims to investigate immunological and virological markers of HIV infection during pregnancy and to inform understanding of the impact of antenatal antiretroviral therapy regimens on such markers, while also assessing the presence of treatment-related adverse pregnancy outcomes

Objectives

1. To model the natural variation of HIV RNA viral load during pregnancy.
2. To examine changes in HIV RNA viral load and CD4 cell counts during pregnancy in women receiving HAART at conception and the impact of HAART and other covariates on these markers.
3. To determine whether the choice of initial HAART regimen in ARV-naive pregnant women is associated with the time to achieving undetectable viral load by the time of delivery.
4. To assess whether ART use during pregnancy is associated with an increased risk of birth defects and prematurity.

2.3 Data source and methods

2.3.1 European Collaborative Study

The ECS is a prospective cohort study in which HIV-infected women are enrolled during pregnancy and their children followed from birth according to the standard ECS protocol (European Collaborative Study 1988;European Collaborative Study 1999;European Collaborative Study 2005b) (Appendix A). The ECS has been ongoing since 1986, and by January 2007, 7176 infected women had been enrolled from 23 centres in 10 countries; Figure 2.1 shows the locations of the principal ECS centres.

General methodology

The ECS consists of an epidemiological coordinating centre at the Institute of Child Health, University College London in London, UK and a network of expert clinicians and researchers (Appendix B). At each ECS centre a dedicated paediatrician or

Figure 2.1 Location of principal centres participating in the ECS

Key	Centre
1	Padua
2	Berlin
3	Edinburgh
4	Madrid
5	Valencia
7	Amsterdam
8	Stockholm
9	Genoa
10	London
11	Brussels
12	Barcelona
13	Copenhagen
15	Milan
16	Turin
17	Naples
18	Warsaw
19	Odessa
20	Micolaiev

obstetrician is responsible for the coordination of enrolment and follow-up, and liaison with the coordinating centre. Standardised data collection forms according to ECS protocol are used in all centres to record routinely collected data, including important socio-demographic details and clinical and laboratory test results; thus no extra patient visits or tests are required as a result of enrolment in the study (Appendix A).

Data collection

ECS centres systematically identify HIV infected pregnant women according to local practice; a well defined policy of routine antenatal testing exists in all centres. Women identified as HIV-infected before, during pregnancy or at delivery are eligible for inclusion. Consenting women are followed up during pregnancy, with their children (both infected and uninfected) followed up prospectively from birth.

Clinical and laboratory information is collected from enrolled women 1-3 times during pregnancy; the amount of follow-up information depending on the trimester at which the women were enrolled and the number of routine visits they underwent during pregnancy. Initial information collected includes mother's obstetric history, marital status, ethnic group, likely route of HIV acquisition (injecting drug use (IDU), needle sharing, high risk sexual partner), HIV test history and current HIV status including treatment details.

During pregnancy and at delivery, virological and immunological data on both qualitative and quantitative HIV DNA PCR and HIV RNA PCR are reported with the assays used given, as well as the results of laboratory investigations to detect total lymphocytes, CD4 and CD8 cell counts.

Required delivery information includes gestational age in completed weeks, birthweight in grams, the administration of prophylactic ART during labour or delivery including type of therapy and dose, mode of delivery and any perinatal problems after delivery, including the presence of congenital abnormalities.

Data management

Follow-up questionnaires from ECS centres are returned to the coordinating centre in London on a regular basis. The data received from the clinicians undergo a number of data checks for completeness before being entered into a Microsoft Access Database which holds all information on child and maternal follow-up since the beginning of the ECS in 1986. In addition to this data collection method, there are several centres that send their data electronically, usually in the form of a Microsoft Excel spreadsheet. Any additions and changes from the previous information provided by that centre are checked by the ECS coordinator before being merged into the ECS database.

Prior to any follow-up data being sent to the coordinating centre, the questionnaires are seen by the key ECS clinician at each centre and so an internal data quality check exists at this initial stage. Data entry in London is carried out by the ECS coordinator, ensuring a high level of consistency and accuracy. Any inconsistencies evident at the time of data entry are queried with the clinician who will return to the patient's medical notes to verify the information.

Additional logical data checks were carried out before carrying out any analyses and included confirmation that:

- Women who were ART-naïve did not have any previous pregnancies recorded in the ECS in which any type of antiretroviral treatment was received

- Dates of ART initiation or HIV RNA and CD4 measurements were not from before the timing of first HIV diagnosis
- Unusual blips in CD4 counts and HIV RNA measurements, and undetectable values of HIV RNA among women who were ART-naïve were corroborated against the original forms to rule out any data entry errors.

Ethical procedures

Approval is in place for the ECS from the coordinating centre, Institute of Child Health/Great Ormond Street Hospital, and conforms to current guidelines on patient confidentiality and anonymity of data collected and published. Each ECS centre obtains local ethical approval updated when necessary according to local protocol. This ensures the correct ethical approval and guidelines are followed, both at the point of data collection and at the point of data analysis and dissemination.

Information from patients enrolled in the ECS is collected at routine follow-up visits and all standard data would be collected by the obstetricians or paediatricians as a matter of course and for this reason, written informed consent is not considered necessary for the collection of this standard data. However, all participants are given details about the project when invited to enrol, informed that the data will be forwarded to the ECS coordinating team anonymously and the option to decline participation or withdraw from the study at any point without detrimental consequences on their care and management.

ECS representative survey

In 2005, a small survey was carried out in 11 of the principal ECS centres (Amsterdam, Barcelona, Berlin, Brussels, Madrid, Milan, Naples, Stockholm, Turin and Warsaw) to quantify the level of non-participation for the years 2002-2004 and the generalisability of

the findings from ECS data. Information requested included the numbers of women declining to enrol and the reasons for non-participation.

Of the 724 women eligible for inclusion from the 11 ECS centres in this period, 4.4% (32) were not enrolled. The principal reason for this was relocation of the patient to another country or area. No centres cited refusal to enrol as a reason for non-participation, therefore limiting any potential bias for the exclusion of these women, i.e. non-enrolment was based on exogenous factors.

2.4.1 HIV RNA viral load

HIV RNA viral load is quantified as the number of copies per millilitre (copies/ml) of blood plasma or serum through the use of quantitative Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) assays, with the type of assay used recorded. The quantification limit of each assay system does not allow detection of viral load below the assay-specific cut-off value and these HIV RNA measurements are therefore left-censored. HIV RNA values are conventionally $\log_{10}$ transformed to resolve non-normality and heteroscedasticity.

In the ECS, a small number of maternal plasma and serum samples for women delivering before 1998, when routine HIV RNA testing was not available in most centres, were frozen locally and shipped in dry ice to either one of two laboratories in Padua and Stockholm where they were stored at -70°C until testing in 1997/1998.

For HIV RNA quantification, the Amplicor HIV-1 Monitor Tests (Standard, version 1.5 with a quantification limit of 400 copies/ml; Ultra-sensitive with a quantification limit of 50 copies/ml: Roche Diagnostic Systems Inc., Branchburg, NJ, USA), Organon- Teknika

(NASBA) assay (Standard NASBA with a quantification limit of 500 copies/ml; Nuclisens with a quantification limit of 40 copies/ml: Organon Teknika, Oss, the Netherlands) or Quantiplex HIV-1 RNA (b-DNA) assays (version 3.0 with a quantification limit of 50 copies/ml: Chiron Diagnostics, Emeryville, CA, USA) were used.

2.4.2 Immunological markers

Standard, locally performed laboratory tests for CD4 cell count and absolute lymphocyte measurements are based on flow cytometry. CD4 cell counts are recorded in units of 10^6 cells/per litre, but following international reporting standards, will be presented throughout this thesis in cells/mm^3.

CD4 cell measurements were square root or $\log_{10}$ transformed to resolve non-normality and heteroscedasticity, as appropriate. CD4 cell counts are also expressed in three categories: severely suppressed immune status was defined as <200 cells/mm^3, moderate immune status as 200-499 cells/mm^3 and competent immune status as ≥500 cells/mm^3.

2.5 General definitions

Black women in the ECS are mostly from countries in sub-Saharan Africa. As most women in the ECS are either Caucasian or black, investigations of race were restricted to comparisons between these two groups.

The first trimester of pregnancy was defined as up to and including 12 gestational weeks, the second trimester as 13-27 gestational weeks and the third trimester as ≥28 gestational weeks.

Premature delivery was defined as delivery before 37 weeks gestation and severe premature delivery as before 34 weeks. Low birth weight was defined as <2500 grams.

2.6 Data analysis

Data entry and management were carried out using Microsoft Access 2003 (Microsoft Corp., Redmond, Washington, USA). Data cleaning and analyses were carried out using Stata version 9.1 (StataCorp LP, College Station, Texas), SAS statistical software version 9 (SAS Institute, Cary, North Carolina, USA) and R version 1.90 -2.4.1 (R Development Core Team 2006).

2.7 Statistical methods

2.7.1 General statistical methods

Continuous variables were compared using the t-test and skewed continuous variables were compared using the Wilcoxon/Mann-Whitney test (Kirkwood and Sterne 2003). Univariate comparisons for categorical variables were tested with the χ^2 test or χ^2 test for linear trend (Kirkwood and Sterne 2003). Fisher's exact test was used to test univariate comparisons of categorical variables when expected numbers were small. The exact method was used to calculate 95% confidence intervals (CI) for binomial proportions.

Univariable and multivariable logistic regression were used to investigate risk factors associated with binary outcomes and to obtain unadjusted and adjusted odds ratios, respectively. Model selection was based on the analysis of deviance between nested models (McCullagh and Nelder 1983).

2.7.2 Methods specific to each section

When more complicated statistical methods have been used (which have not been covered above), detailed descriptions are given in each corresponding results chapter. A summary is given below:

Chapter 3– Univariate linear mixed effects model with nested random effects and an additional variance function to model heteroscedasticity (Pinheiro and Bates 2000).

Chapter 4 – Turnbull non-parametric estimator, interval censored parametric survival model (Turnbull 1976;Powderly *et al.* 1999;Collett 1994), propensity score adjustment (D'Agostino, Jr. 1998).

Chapter 5 – Univariate linear mixed effects model accounting for left-censored viral load measurement using a conditional likelihood approach (Thiébaut and Jacqmin-Gadda 2004), bivariate linear mixed effects models dealing with left-censored measurements with imputed midpoints (Thiébaut *et al.* 2002) and a conditional likelihood approach (Thiébaut and Jacqmin-Gadda 2004;Thiébaut *et al.* 2005), local dependence map (Jones and Koch 2003)

Chapter 6 – Weighted Cox regression model (Therneau and Grambsch 2000)

2.8 Data available for each analysis

By December 2006, the number of women enrolled in the ECS was 7,176. The analyses presented here focused on a specific sub-group of women in the ECS and for these reasons, the total number of women in each analysis varied; Table 2.1 shows the numbers of women/mother-child pairs used in each analysis.

Table 2.1 Numbers used in analyses

Analysis	Chapter	No. of Women	Time of analysis
HIV RNA viral load in untreated women in pregnancy	3	333	2005/2006
Time to undetectable HIV RNA viral load	4	240	2006
Joint model for CD4 count and HIV RNA viral load in pregnancy	5	162	2007
Risk of congenital abnormalities	6.1	3740	2004
Risk of prematurity	6.2	1294	2005/2006

2.9 Role of researcher

As this work was carried out as part of the ongoing study, which I joined in 2004 as the ECS statistician, it is appropriate that my role in the research presented in this thesis is clarified. The clinicians within each centre were responsible for data collection, data verification, entry and database management was carried out by the ECS coordinator at the coordinating centre at the Institute of Child Health. I was given the responsibility for the preparation and analysis of the data relating to HIV-infected women in particular, with the aim to inform the understanding of HIV infection during the period of pregnancy.

The direction and nature of the analyses were agreed in consultation with my supervisor Professor Marie-Louise Newell and co-supervisors Drs. Mario Cortina-Borja and Claire Thorne. The strategic and technical development of the analyses presented in this thesis was entirely my responsibility. The work has resulted in two manuscripts for publication in peer-reviewed journals (Appendix C); I took the lead role in preparation of these

manuscripts, as part of the ECS team. The final rates of congenital abnormalities by type and timing of ART regimen were also sent to the Antiretroviral Pregnancy Registry (Antiretroviral Pregnancy Registry Steering Committee 2004) and were published as part of their interim report.

Chapter 3 Determinants and trends of HIV RNA levels in ART-naïve pregnant women

3.1 Introduction

Data on whether increasing gestation may impact on HIV RNA levels - the pre-eminent risk factor for MTCT - are limited and mostly suggest that levels remain stable during pregnancy, increasing in the early postpartum period (Burns *et al.* 1998;Mayaux *et al.* 1997a;Cao *et al.* 1997;O'Shea *et al.* 1998). However these studies included only small numbers of untreated women, or included women who were mostly on monotherapy making it difficult to separate a possible gestation effect from that of a treatment effect. Additionally these studies either did not adjust for confounders in adjusted analyses, such as maternal CD4 counts or did not use appropriate methods for repeated measures data with the analyses mostly limited to the third trimester of pregnancy.

This analysis used pregnancy data on over 300 untreated HIV-infected women enrolled in the ECS with HIV RNA measurements from the second trimester of pregnancy, a comparatively larger number of exclusively untreated women than other published studies over a longer period of pregnancy. Trends in HIV RNA viral load over pregnancy were examined using a LME model, an appropriate statistical method and adjusted for factors known to be associated with the level of HIV RNA level, including CD4 cell counts.

3.2 Methods specific to this chapter

3.2.1 Data

This analysis was restricted to the 333 women who were reported as untreated prior to and throughout pregnancy and who had at least one HIV RNA measurement during pregnancy. Women receiving intrapartum only zidovudine for prevention of MTCT were not excluded; however women receiving any antenatal ARV prophylaxis were. As women delivered at different gestational ages, a similar approach to that adopted by Watts and colleagues (Watts *et al.* 2003) was used with HIV RNA viral load modelled with respect to the time from delivery. This approach avoids the grouping of measurements by trimester of pregnancy and makes full use of the data; therefore time from delivery ranged from -35 weeks to 0 weeks, with 0 weeks representing the time of delivery for all women. To improve the regression estimates and to examine interactions between ethnicity over pregnancy (only white women had measurements in early pregnancy), the few measurements available between -35 and -25 weeks (roughly corresponding to 5 and 15 weeks gestation, respectively) were excluded; this equated to the removal of 23 measurements and 9 women (with only one measurement in pregnancy) from the regression analysis.

HIV disease progression is associated with the time from seroconversion (Collaborative Group on AIDS Incubation and HIV Survival including the CASCADE EU Concerted Action 2000) and HIV RNA levels levels have been shown to vary by CD4 cell counts (Mellors *et al.* 1997), assay type (European Collaborative Study 2002), plasma or serum used (European Collaborative Study 1999), race and history of injecting drug use (Anastos *et al.* 2000;Saul *et al.* 2001;Touloumi *et al.* 2004). The following variables, together with time from delivery (weeks), were therefore considered in regression models; race (white or black), first CD4 cell count measured in pregnancy (≥500 200-

499 or <200 cells/mm^3), type of HIV RNA quantitative assay used (NASBA/Nuclisens or Roche), material used (plasma or serum), timing of first HIV diagnosis (before pregnancy or during) and history of IDU (Y/N).

3.2.2 Statistical methods

Modelling HIV RNA measurements over time is complicated by the repeated measures and left-censored nature of this type of longitudinal data, as well as the potential non-linear pattern of HIV RNA over pregnancy. The vast majority of HIV RNA viral load measurements were from women enrolled before 1998, when viral load measurements were not routinely collected (Section 2.4.1). These earlier measurements were stored for longer and the extent to which this may impact the accurate quantification of HIV RNA is not known; therefore it is prudent to attempt to account for any resulting variation between samples in the regression model.

Linear mixed effects model

The repeated measures nature of the data, obtained from the two sampling periods (i.e. the periods before and during routine collection) require a model which can account for any variation within and between women and also any variation between the respective sampling periods. A standard linear mixed effects (LME) model (Pinheiro and Bates 2000;Goldstein 1989) with nested grouping factors, otherwise known as a multilevel model, can be used to analyse these data by incorporating a random effect on the intercept and slope of HIV RNA viral load over gestation, according to the sampling period.

A multilevel LME model for the level of HIV RNA viral load y_{ijk} at the kth week of pregnancy of the jth woman within the ith sampling period is

$$y_{ijk} = \left(\beta_0 + b_{0i} + b_{0ij}\right) + \left(\beta_1 + b_{1i} + b_{1ij}\right)k + \varepsilon_{ijk}, \quad i = 1,...,M;\ j = 1,...,M_i;\ k = 1,...,K$$

$$\mathbf{b}_i = \begin{bmatrix} b_{0i} \\ b_{1i} \end{bmatrix} \sim N\left(\mathbf{0}, \mathbf{\Psi}_1\right), \mathbf{b}_{ij} = \begin{bmatrix} b_{0ij} \\ b_{1ij} \end{bmatrix} \sim N\left(\mathbf{0}, \mathbf{\Psi}_2\right)$$

$$\varepsilon_{ijk} \sim \mathrm{N}\left(0, \sigma^2\mathbf{I}\right).$$

where M is the number of first level groups (two in this case, before and after routine collection of viral load) and M_i is the number of second level groups within the first level group i (in this case the number of women within each sampling period). The parameters β_0 and β_1 are the fixed effects for the intercept and slope in the model. The random effects b_{0i} and b_{1i}, corresponding to the intercept and slope for the sampling period are assumed to be independent for different $i$with h mean zero and variance-covariance of Ψ_1. The women-within-sampling period random effects for intercept and slope, b_{0ij} and b_{1ij}, are assumed to be independent for different i and j and to be independent of b_{0i} and b_{1i} with mean zero and variance-covariance Ψ_2, where Ψ_1 and Ψ_2 are symmetric positive-definite matrices from a multivariate Gaussian distribution. The within-group errors ε_{ijk} are assumed to be independent for different i, j and k with mean zero and variance $\sigma^2\mathbf{I}$, and independent of the random effects. Using this multilevel approach, the variance components associated with the different levels of nesting and the within group error can be estimated.

All women enrolled in the ECS will be seen regularly in pregnancy as part of their routine antenatal care, but this will not always include testing for viral load or CD4 counts. This analysis used all available data for a given woman, and because the timing of CD4 cell and HIV RNA testing was dictated by clinical indication instead of a predetermined schedule, there are no missing data in the classic sense in these measurements. Further, the LME model assumes that the different women in the model

are the only completely independent units of observation in the study and that viral load measurements within each woman are representative of that woman's viral load trajectory in pregnancy (Leyland and Goldstein 2001;Goldstein 1989). Assuming that these assumptions hold, there is no restriction in the number or spacing of measurements in an LME model, allowing for a single model to be fit to individuals who have one or several measurements (Leyland and Goldstein 2001). Therefore in the LME models presented here, women with only one HIV RNA measurement were included; however, they will contribute more to the estimate of the intercept of the model than the slope (Goldstein 1989).

Advanced LME modelling techniques which have been developed to account for left-censoring of HIV RNA viral load are not able to incorporate a nested random effects structure at present (Thiébaut *et al.* 2005). For this reason, a standard multilevel LME model was used to study changes in HIV RNA in pregnancy, with midpoints imputed for measurements which were recorded at the assay lower limit of quantification.

Scatterplots of $\log_{10}$ transformed viral load measurements over pregnancy were used to examine the data. A supersmoother, a non-parametric (data-driven) smoothing function capable of summarising the trend of a response variable, y, as a function of a continuous covariate, x, provided a smooth estimate of the changes in overall viral load measurements over pregnancy (Friedman 1984). The supersmoother is based on a symmetric k-nearest neighbour linear least squares procedure, which involves taking a window of size $k/2$ data points on each side of x and fitting a regression line to that point to predict the value of x that best describes the points in that window. This local regression is run for three values of k, $n/2$, $n/5$ and $n/10$, and the value of k that leads to the best estimate for each window is determined using cross-validation (Venables and

Ripley 1997;Friedman 1984). The supersmoother is preferred here over a Kernel or lowess smoother because it is adaptive to the amount of smoothing that is required.

Polynomials and spline smoothing

LME models with orthogonal polynomials and smoothing splines in the fixed effects were compared in modelling the basic pattern of HIV RNA viral load on gestational weeks to delivery.

Orthogonal polynomials with linear, quadratic and cubic terms were used to avoid collinear terms resulting from use of conventional (non-orthogonal) polynomials. Spline smoothing of the covariate is a more appropriate approach when the association between a continuous covariate and the outcome changes in a more complicated way, i.e. one which cannot be easily summarised by low-order polynomials, and is explained briefly in the following section.

A LME model with a smoothing spline is able to describe the local structure of the relationship between the outcome and the covariate providing a good fit to the data across the full range of the covariate and is described below (Gurrin *et al.* 2005;Venables and Ripley 1997).

Suppose that the relationship between a continuous response Y and a single continuous covariate x is to be modelled by

$$\mathrm{E}(Y_i) = m(x_i) + \varepsilon_i, \quad i = 1,...,n$$

where m is an arbitrary smooth function giving the conditional mean of Y, ϵ_1, ϵ_2, ... ϵ_n are independent, normally distributed random error terms with common mean zero and

variance σ_ε^2. This represents a non-parametric regression model where *m* can take on any form; splines are a popular method used for representing *m*, and are comprised of piecewise polynomials (usually linear, quadratic or cubic) joined together at points known as knots.

The linear spline estimator is of the form

$$m(x) = \beta_0 + \beta_1 x + \sum_{k=1}^{K} u_k (x - \kappa_k)_+$$

where

$$(x - \kappa_k)_+ = \begin{cases} 0, & x \le \kappa_k \\ x - \kappa_k, & x > \kappa_k \end{cases}$$

and $\kappa_1, \kappa_2, \ldots, \kappa_K$ are knots. The first of the above equations describes a sequence of line segments tied together at the knots to form a continuous function. This can be extended to a piecewise cubic polynomial given below, defined as

$$m_3(x; \boldsymbol{\beta}, \mathbf{u}) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \sum_{k=1}^{K} u_k (x - \kappa_k)^3_+$$

where $\boldsymbol{\beta} = (\beta_0 + \beta_1 + \beta_2 + \beta_3)^T$ and $\mathbf{u} = (u_1, \ldots u_k)^T$ denote vectors of coefficients. The cubic spline with the given knots is then a linear combination of the basis functions $1, x, \ldots, x^p, (x - \kappa_1)^p_+, \ldots, (x - \kappa_K)^p_+$ and is appropriate for smoothing in a LME model. A cubic spline was preferred here as it results in a smoother fitted curve than the linear or quadratic spline (Gurrin *et al.* 2005).

3.2.3 Methods used for comparing and assessing models

Likelihood ratio tests

The likelihood ratio test (LRT) is a general method for comparing nested models fit by maximum likelihood such as the LME models described above. A statistical model is

said to be nested within another model if it represents a special case of the other model. For example if model 1 includes a linear term for gestational age only and model 2 includes the same linear term and an additional binary covariate for race, then model 1 is nested within model 2, as model 1 is a special case of when the coefficient for the binary covariate for race is zero. If L_2 is the likelihood of the larger model (model 2) and L_1 is the likelihood of the simpler model (model 1) then we would have $L_2 > L_1$, and $\log L_2 > \log L_1$. The LRT statistic is then given by

$$\Lambda = 2\log(L_2 / L_1) = 2\left[\log(L_2) - \log(L_1)\right].$$

If k_i is the number of parameters estimated in model i, then the asymptotic distribution of the LRT under the null hypothesis that the smaller model is adequate (i.e. adding additional terms does not improve the fit of the model) is a χ^2 distribution with k_i-k_{i-1} degrees of freedom (d.f.); the LRT statistic and the associated p value from the $\chi^2_{k_i - k_{i-1}}$ distribution will be used here to assess improvements between nested models from including additional random effects terms and their covariates.

Akaike Information Criteria

The Akaike Information Criteria (AIC) (Akaike 1974) is a model comparison criterion evaluated as

$$AIC = -2\log \text{Lik} + 2p,$$

where p is the number of parameters in the model, and is a useful tool in examining two models which are not nested but which are based on the same dataset (Pinheiro and Bates 2000;Akaike 1974). When comparing two models, the model with the smaller of the two AICs is the preferred model.

A combination of LRTs and AIC comparisons were used to guide the model selection procedure. Standard errors and Wald tests, which inform the statistical significance of specific terms within a model, were also examined to assess individual covariates in fitted models.

3.2.4 Model selection procedure

In order to establish the appropriate functional form of HIV RNA changes over pregnancy, the orthogonal polynomial and smoothing cubic spline LME models were first evaluated. After determining the appropriate functional form between HIV RNA levels and time from delivery, the inclusion of additional potential confounders was tested.

The covariates listed earlier in this methods section were then considered alone in the model with time from delivery to assess whether their inclusion resulted in an improved log-likelihood (LRT~ χ^2_{df} ; $p \leq 0.10$). Covariates which improved the fit of the model were then included together in a subsequent model (Collett 1994). Any covariates which were not included at this stage were further tested for inclusion in the model with the LRT, as the presence of the other covariates in the model may now necessitate their addition. Additionally, to examine whether the fit of the model could be improved by removing covariates already in the model, the LRT from removal of each of these terms in turn (while holding the other covariates in the model constant) was assessed. The initial model was chosen when no improvements in the log-likelihood could be attained through the addition or removal of any main effects.

Main effects as explanatory factors in the initial model imply that their effect on HIV RNA is constant over gestation and with respect to other covariates in the model; the

inclusion of interaction terms test the joint effects of gestational age and other covariates. Covariates in the model were therefore assessed for significance of interactions with gestational age, and where appropriate with one another, and were included if inclusion resulted in a significant improvement in the log-likelihood. To ensure that all important main effects were included, a conservative p-value of 0.10 associated with the likelihood ratio test was used, however a less tolerant $p<0.05$ was used for inclusion of interaction terms. The presence of an interaction term requires the inclusion of the main effects terms, regardless of the significance of the associated Wald test (Cox *et al.* 1984).

The LME model was carried out using the *lme* function in R version 2.2.0 (R Development Core Team 2006).

3.3 Results

The characteristics of the 333 pregnant women in this analysis are given in Table 3.1, along with the number of actual HIV RNA measurements available by characteristic (e.g. the number of measurements on women which were quantified with a NASBA assay or from blood plasma). A fifth of women were black, with the majority born in sub-Saharan Africa. There were no differences in CD4 cell counts between black and white women (median 460 [IQR 20-657] cells/mm^3 vs. 460 [IQR 70-550] cells/mm^3; Wilcoxon test $p=0.18$), but black women were more likely to have had their first HIV diagnosis during pregnancy (45/66 vs. 93/247; $\chi_1^2=15.5, p<0.0001$).

Over a third (35%) of women had at least two HIV RNA measurements, with a median of three measurements per woman (IQR 3-4) and the remainder of women had only one measurement available in pregnancy. Most women with only one HIV RNA measurement available in pregnancy (173/218; 79%) had this taken at delivery. Women

Table 3.1 Baseline characteristics for women and measurements

Characteristic	Women†	HIV RNA measurements on women†
	n=333	*n*=490
Viral load values (copies/ml)		
Median (IQR)	-	2000 (IQR 350-11,000)
Viral load measurement		
Above assay detection level	-	338 (69)
Below assay detection level	-	152 (31)
First CD4 cell count (cells/mm^3)		
Median (IQR)	437 (IQR 287-633)	-
		-
Timing of CD4 measurement (weeks from delivery)	-13 (IQR -19,-3)	-
Missing	45 (14)	
Race		
White	249 (77)	356 (76)
Black	66 (21)	101 (21)
Asian	6 (2)	14 (3)
Unknown	12	19
Age at delivery (years)		
Median (IQR)	28 (25-31)	
15-26	108 (33)	158 (33)
26-31	140 (43)	198 (41)
≥32	78 (24)	127 (26)
Unknown	7	7
History of IDU		
Non-IDU	156 (49)	248 (52)
IDU	162 (51)	224 (48)
Unknown	15	18
Timing of HIV diagnosis		
Pre-pregnancy	182 (55)	287 (59)
During pregnancy	151 (45)	203 (41)
HIV RNA assay		
NASBA/Nuclisens	-	188 (38)
Roche	-	283 (58)
Other	-	19 (4)

† - *n* (%) unless otherwise stated

Table 3.1 contd.

Characteristic	Women†	HIV RNA measurements on women†
Sampling period		
Routine collection	-	98 (20)
Before routine collection	-	392 (80)
Blood material		
Plasma	-	286 (59)
Serum	-	200 (41)
Time period of delivery		
1987-1993	228 (68)	350 (71)
1994-1998	40 (12)	52 (11)
1998-2005	65 (20)	88 (18)

† - *n* (%) unless otherwise stated

with only one HIV RNA measurement had similar CD4 cell counts to women with at least two measurements (median 433 cells/mm^3 [IQR 285-647] vs. median 450 cells/mm^3 [IQR 289-610]; Wilcoxon test p=0.9), but had significantly lower levels of HIV RNA at delivery (median 2.30 $\log_{10}$ copies/ml [IQR 2.00-3.55] vs. median 3.49 $\log_{10}$ copies/ml [IQR 3.04-4.17]; Wilcoxon test p<0.001).

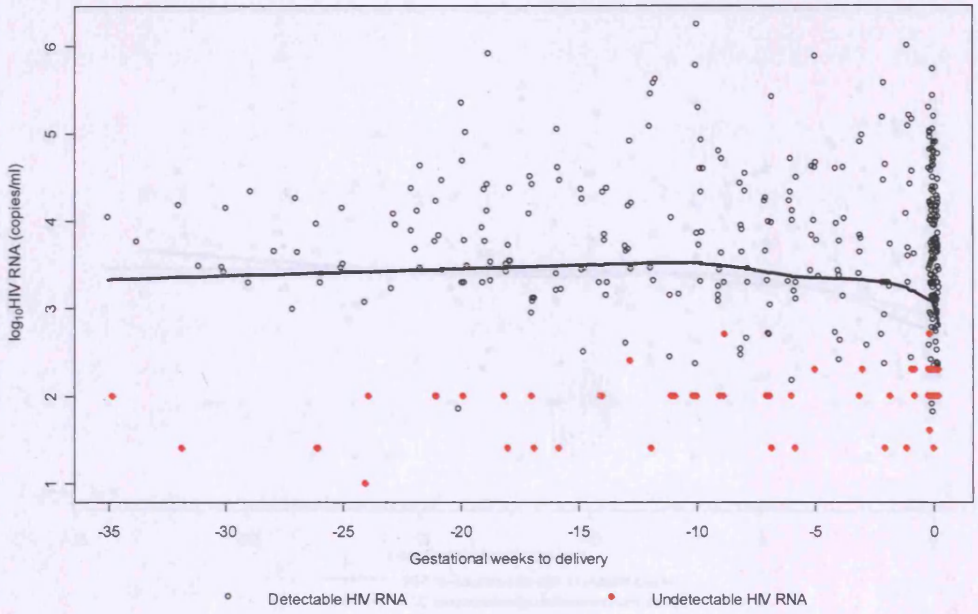
Nineteen women had HIV RNA measurements taken in the first trimester, 83 in the second trimester, 120 in the third trimester and 241 at delivery; the median HIV RNA levels were, respectively for these times, 3.50 (IQR 3.35-4.01), 3.51 (IQR 3.10-4.20), 3.52 (IQR 2.70-4.34) and 2.76 (IQR 2.00-3.75) $\log_{10}$ copies/ml. The overall proportion of undetectable values was surprisingly high (31%) in this untreated group of women. Of the 152 undetectable values of HIV RNA, 13 (9%) were <50 copies/ml, 91 (60%) were <200 copies/ml, 45 (29%) were <400 copies/ml and the remaining 3 (2%) were measured as undetectable with assays with detection limits of <500 and <1000 copies/ml; the majority (131/152; 86%) of the undetectable measurements were from

blood samples taken before 1996. HIV RNA levels were measured in plasma for 59% (n=286) of samples and were significantly lower than those measured in serum (median 3.15 [IQR 2.30-3.95] vs. median 3.48 [3.00-4.19]; Wilcoxon test p<0.001).

3.3.1 HIV RNA viral load over gestational age

A supersmoother was used to reveal the structure of the $\log_{10}$ transformed HIV RNA viral load data with respect to time from delivery and is shown in Figure 3.1, with detectable and undetectable measurements marked. A jitter of the time of HIV RNA measurement was added to the plot to elucidate the number of measurements at delivery and to aid smoothing of the curve over pregnancy. A slight linear increase in HIV RNA levels over pregnancy can be observed until -13 weeks, followed by a slight linear decrease until -5 weeks at which point viral load is seen to decrease more rapidly up to the time of delivery. The small numbers in the first trimester of pregnancy (approximately corresponding to measurements less than -25 weeks in the scatterplot) were apparent. Subsequent models excluded these early measurements and were based on 467 measurements and 324 women.

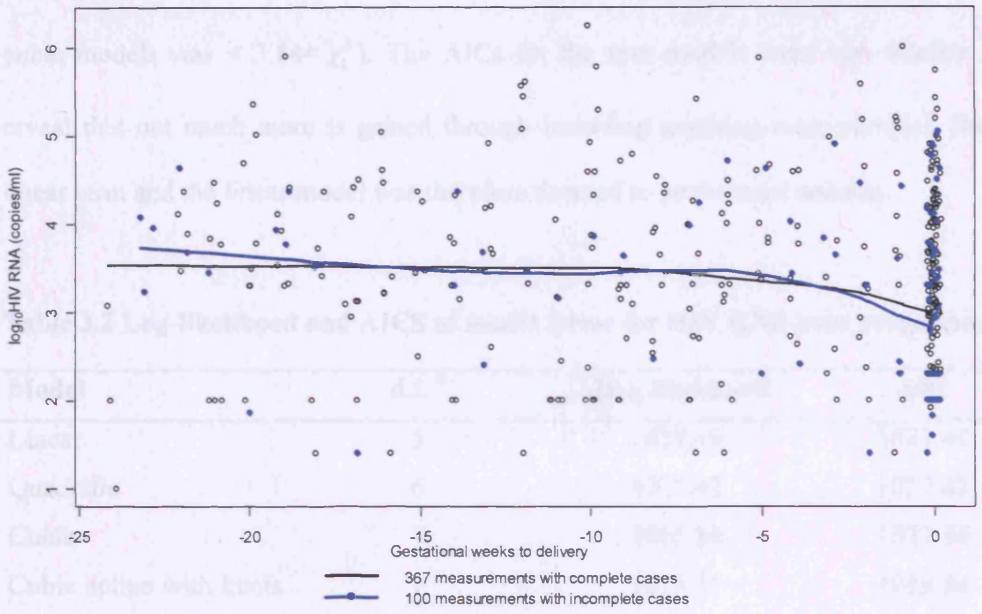
Figure 3.1 Scatterplot of HIV RNA viral load over pregnancy with supersmoother



3.3.2 Model selection procedure

For 86 of the 467 HIV RNA measurements available, information was missing on at least one of the variables considered for inclusion in the model; most of these were missing information on CD4 counts (45/86) and race (19/86). In order to test for an interaction term between race and time from delivery over the 24 weeks, the 14 measurements on Asian women were also removed from the analysis. The remaining 367 measurements were used in subsequent modelling. The measurements included in regression models were overall similar to the 100 measurements not included with respect to the distribution of HIV RNA (median 3.30 [IQR 2.56-4.08] vs. 3.30 [IQR 2.43-3.94] $\log_{10}$ copies/ml; Wilcoxon test, p=0.41), the proportion of censored values (110/367 vs. 38/100; χ^2_1=1.84, p=0.17) and the crude pattern of HIV RNA with respect to gestational weeks from delivery (Figure 3.2).

Figure 3.2 Scatterplot of HIV RNA viral load over pregnancy with a supersmoother by HIV RNA measurements included in the model and those that were not



The simplest model, a LME model with a linear term in weeks from delivery was fitted first, with a separate random intercept term for the sampling period. The inclusion of a random effect term for the slope of weeks from delivery did not lead to a significant improvement in the model (LRT $\chi_1^2 \approx 0$; $p \approx 1$) and the associated estimated standard deviation for the random effect term for the slope was extremely small. This suggests the presence of a substantial amount of variation of HIV RNA among the intercepts for individual women, but not among their slopes for gestational weeks from delivery and the random effects term for the slope is not required.

An orthogonal quadratic and cubic model for gestational weeks from delivery (GA) with the same random effects terms was also fitted in a LME model and the associated AICs and log-likelihoods are given in Table 3.2. The supersmoother in Figure 3.1 revealed fluctuations in HIV RNA levels at around -13 and -5 weeks and therefore an LME model with a natural cubic spline was fitted, with knots at these two time points. Adding

additional power terms to the linear and quadratic model, respectively, did not improve the log likelihood of the model (LRT between linear and quadratic, and quadratic and cubic models was $< 3.84 = \chi_1^2$). The AICs for the four models were very similar and reveal that not much more is gained through including anything more complex than a linear term and the linear model was therefore deemed to be the most suitable.

Table 3.2 Log-likelihood and AICS of model forms for HIV RNA over pregnancy

Model	d.f. [a]	-2log likelihood	AIC
Linear	5	1017.49	1027.49
Quadratic	6	1015.42	1027.42
Cubic	7	1015.38	1029.38
Cubic spline with knots at -13 and -5 weeks	7	1015.31	1029.44

[a] – degrees of freedom

The inclusion of each of the main effects of race, first available CD4 cell count in pregnancy, blood material and IDU to the model with a linear term for gestational weeks from delivery only, all led to significant improvements in the log-likelihood. These covariates were then considered together in one model (Table 3.3, model 1).

The inclusion of assay type in model 1 was now associated with a significant improvement in the model (model 2). The model could not be improved further with the addition of other main effects or with the removal of any main effects already in the model. Interactions between covariates in the model were then considered and the associated likelihood ratio tests are also given in Table 3.3. The presence of an interaction term between gestational weeks from delivery and race, and between assay and blood sample type improved the model fit significantly. However, an interaction

Table 3.3 Likelihood ratio tests for model selection procedure

Model no.	Model	LRT	d.f.[a]	*p*
1	Weeks from delivery +race+CD4+IDU+blood type	-	10	-
2	Model 1 + assay type	7.10	1	0.008
3	Model 2 + race* weeks from delivery	7.12	1	0.008
4	Model 3 + assay*blood type	14.84	1	<0.001
5	Model 4 + CD4*race	0.35	2	0.84
6	Model 4 + CD4*IDU	1.41	2	0.50
7	Model 4 + IDU* weeks from delivery	1.73	1	0.19
8	Model 4 + race*blood	2.95	1	0.09
9	Model 4 + race*assay	2.77	1	0.10

[a] - degrees of freedom

term between gestational weeks from delivery and CD4 measurements in pregnancy was not tested as these measurements were not taken at the same time for each woman, while an interaction between IDU and ethnicity could not be tested as there were no black IDUs. There was no evidence of interactions between any other covariates included in the model (Table 3.3).

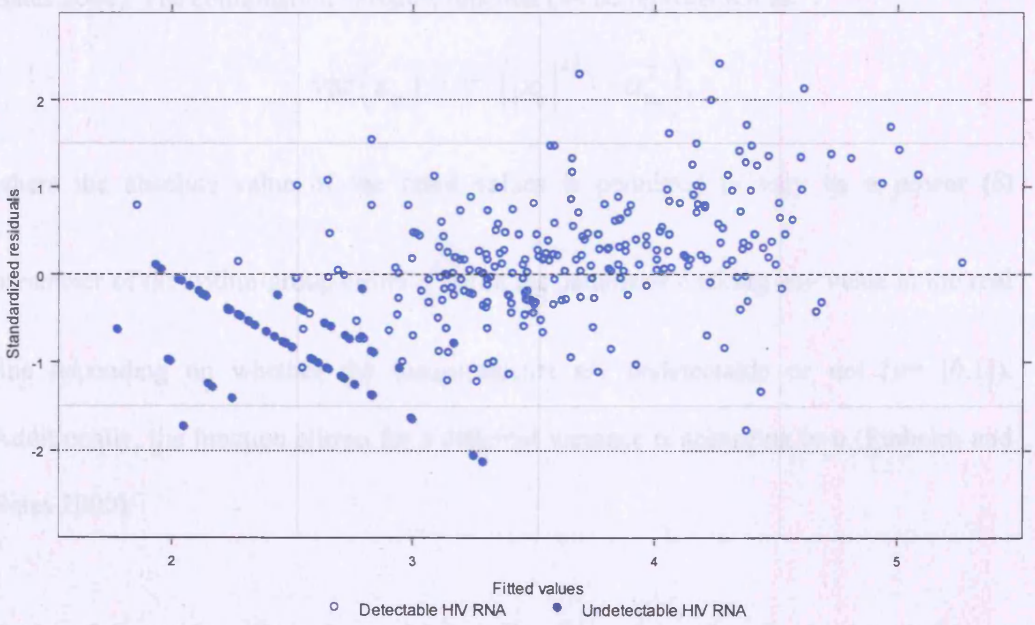
The final model included a linear term in gestational weeks from delivery, CD4 cell count, race, history of IDU, blood type, assay type and an interaction term between gestational weeks from delivery and race, and between blood and assay type. With the exception of IDU, all main effects or their related interaction terms in the model also had significant Wald tests. The removal of the IDU term from the model was not associated with a significant increase in log-likelihood (LRT χ_1^2=2.24; p=0.13) and could on these grounds be removed from the model. However, almost two-thirds of white women were

IDUs and excluding this factor led to a substantial increase in the estimate of the model intercept and the race effect (with the standard errors remaining relatively stable) and the IDU term was therefore retained in the final model.

3.3.3 Model residuals

A primary tool for investigating heteroscedasticity is a plot of the fixed effects standardised (Pearson) residuals versus the fitted values from this model and is shown in Figure 3.3, with residuals from undetectable HIV RNA measurements indicated. The plot gives some indication of within-group residual error changing with fitted values of HIV RNA and reveals a linear pattern in the undetectable residuals. This suggests lack of homoscedasticity, one of the assumptions required for LME models.

Figure 3.3 Plot of the standardised residuals versus fitted values for model without variance function



In favour of accounting for the variance associated with each nested grouping in this model, left-censoring of undetectable HIV RNA measurements were dealt with by imputing midpoints. The residual plot indicated that the fitted values for undetectable

values were overestimated and linearly related to the standardised residuals which highlight one of the disadvantages of adopting this method. In an attempt to correct for this within-group heteoscedasticity, a combination variance function (Pinheiro and Bates 2000) was used to model

a) the variance structure of the within-group errors for undetectable values over the range of fitted values, and
b) the residual error variance (as a constant term) for both undetectable and detectable values.

A formal test for the lack of homoscedasticity could then be carried out by examining the associated improvement in the log-likelihood from this model, with a significant improvement implying the need for the variance function to achieve homoscedasticiy of the within-group errors. A power function to model the variance structure is a common choice for modelling the monotonic heteroscedasticity seen in Figure 3.3 (Pinheiro and Bates 2000). The combination variance function can be represented as:

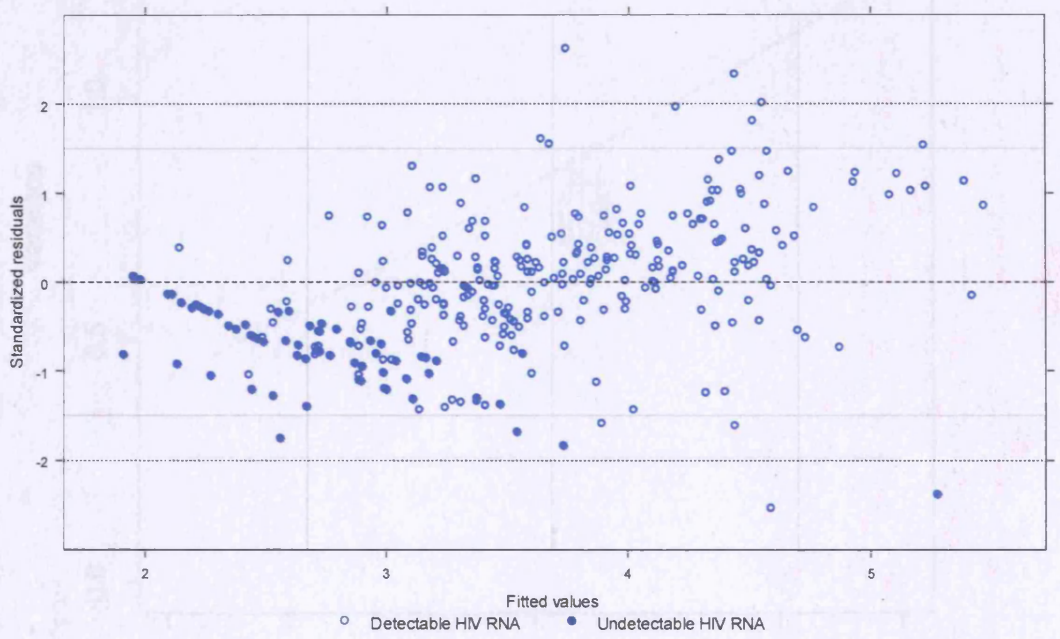
$$\mathrm{var}\left(\varepsilon_{ij}\right) = \sigma^2 \left(\left| x_{ij} \right|^{2\delta_{u_{ij}}} \cdot \alpha^2_{u_{ij}} \right),$$

where the absolute value of the fitted values is permitted to vary by a power (δ) parameter of the within-group errors σ^2, with the parameter δ taking any value in the real line depending on whether the measurements are undetectable or not (u= [0,1]). Additionally, the function allows for a different variance α according to u (Pinheiro and Bates 2000).

A plot of the residuals from the model including the variance function is given in Figure 3.4 and reveals the adequacy of the variance function in representing some of the heteroscedasticity seen in Figure 3.3. The significant improvement in the log-likelihood

(LRT χ^2_2 = 22.95; p<0.0001) and AIC of the model (1027.49 compared to 961.51 for the model without the variance function) confirm the need for the variance function.

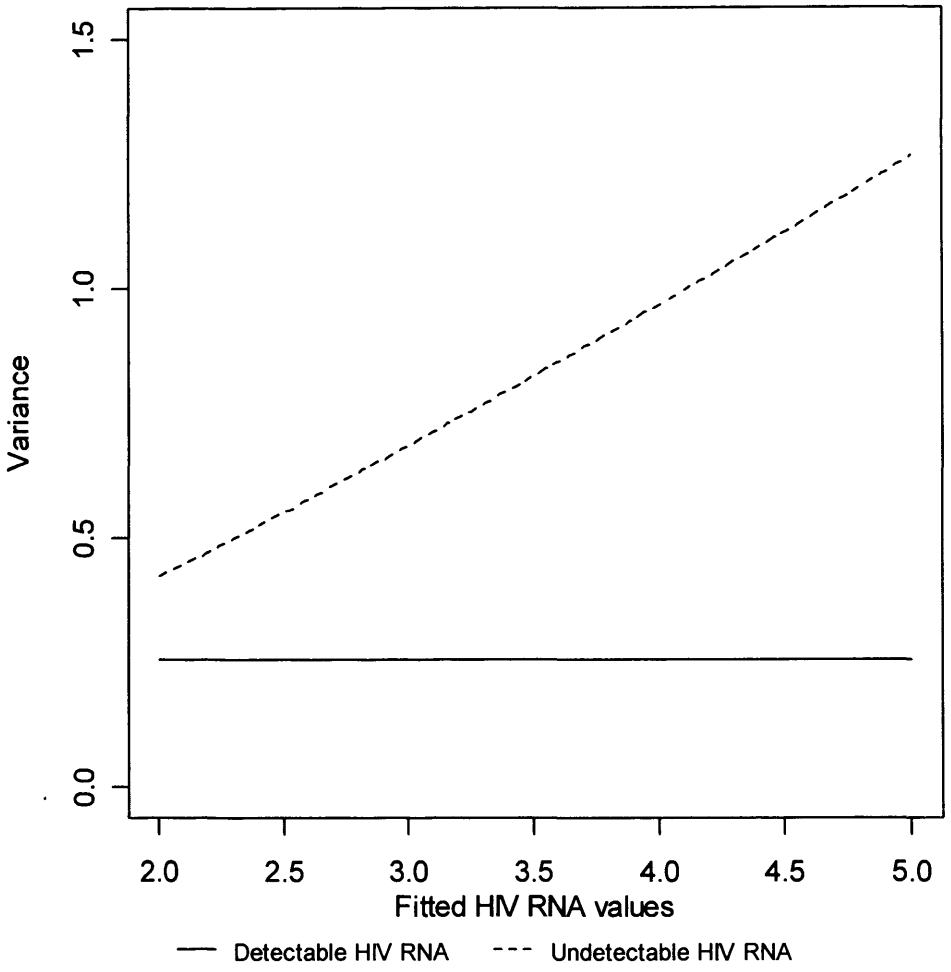
Figure 3.4 Plot of the standardised residuals versus fitted values for model with a variance function to model within-error heteroscedasticity



Estimates for δ and α were: $\hat{\delta} = \begin{cases} 0.0 \text{ if detectable} \\ 0.60 \text{ if undetectable} \end{cases}$ and $\hat{\alpha} = \begin{cases} 1.18 \text{ if detectable} \\ 1.00 \text{ if undetectable} \end{cases}$.

This combination variance function increases with the size of the fitted values for the within-group errors for undetectable measurements only, and remains constant for the within-group errors for detectable values, but is 18% larger than before (Figure 3.5).

Figure 3.5 Estimated variance functions over gestational age, stratified by detectable measurements



3.3.4 Model interpretation

The estimated coefficients of change for mean HIV RNA and the standard deviations for the random effects parameters from this model are presented in Table 3.4. There was a significant association between CD4 cell counts in pregnancy and HIV RNA level; although this should not be interpreted as evidence of causality, women with lower CD4 cell counts were estimated to have higher levels of HIV RNA. HIV RNA levels were also dependent on both the type of assay used and the blood sample type; viral load

Table 3.4 Adjusted coefficients of change for log_{10} HIV RNA viral load in pregnancy and estimated random effects

	Multivariate Coefficient	95% CI	*p* value
HIV RNA at delivery	3.41	2.70, 4.12	<0.001
Gestational weeks to delivery increase	-0.019	-0.03, -0.007	0.002
Race			
White	0.00		
Black	0.45	0.08, 0.83	0.02
HIV RNA assay			
NASBA/Nuclisens	0.00		
Roche	-0.87	-1.22, -0.52	<0.001
CD4 count (cells/mm^3)			
≥500	0.00		
200-499	0.29	0.05, 0.53	0.02
<200	0.78	0.44, 1.11	<0.001
History of IDU			
No	0.00		
Yes	-0.20	-0.47, 0.06	0.14
Sample type			
Plasma	0.00		
Serum	0.20	-0.17, 0.57	0.29
Sample and assay interaction			
Plasma and NASBA/Nuclisens	0.00		
Serum and Roche	1.09	0.60, 1.58	<0.001
Gestational age and race interaction			
Change by GA for white women	0.00		
Change by GA for black women	0.035	0.01, 0.06	0.005
Random effects parameters			
σ intercept between sampling periods (b_{0i})	0.44	0.15, 1.29	-
σ intercept between women (b_{0ij})	0.69	0.60, 0.80	-

copies/ml obtained from a serum sample with a NASBA/Nuclisens assay were 60% (0.20 $\log_{10}$ copies/ml) higher than those obtained from a plasma sample with the same assay, while HIV RNA in a serum sample measured with a Roche assay was estimated to be around 2.5-fold (0.42 $\log_{10}$ copies/ml) higher than that measured in plasma and a NASBA/Nuclisens assay.

From around the start of the second trimester, HIV RNA viral load over pregnancy was estimated to decrease by 0.019 $\log_{10}$ copies/ml per week up to the time of delivery for white women. This corresponds to a 4.3% weekly decrease in copies/ml over pregnancy (95%CI -6.7%, -1.6%), an approximate 65% decrease in HIV RNA levels over the last 24 weeks of gestation. The slope of the change in viral load over pregnancy for black women was positive and significantly different to that for white women as indicated by the *p*-value for the race and gestational weeks from delivery interaction term (Table 3.4). The change in HIV RNA viral load over pregnancy for black women was estimated to be 0.016 $\log_{10}$ copies/ml (95%CI -0.005, 0.037), which equates to a weekly increase of 3.8% copies/ml (95%CI -1.1%, 8.9%), however this was not a significant increase indicated by the zero-bounded 95%CIs.

As the measure of gestation in this model was weeks to delivery, the intercept estimate corresponds to the mean HIV RNA level at delivery for white women, with the race estimate in Table 3.4 corresponding to the difference in mean levels at delivery between black and white women. The estimated mean HIV RNA level at delivery was 3.41 $\log_{10}$ copies/ml (95%CI 2.70–3.41) for white women and 3.86 $\log_{10}$ copies/ml (95%CI 3.16-4.56) for black women.

The estimated standard deviations for the intercept's random effects were 0.44 $\log_{10}$ units for the sampling period and 0.69 $\log_{10}$ units for the between-women intercept; both estimates had 95%CIs which were not bounded by zero. The magnitude of these estimates reveals the considerable variability in the estimated fixed effects mean intercept between women and also between sampling periods (mean HIV RNA at delivery 3.41 $\log_{10}$ copies/ml, Table 3.4). The random effect for the between-women intercept was over 1.5 times larger than that for the sampling period and suggests that individual variability explains more of the variation around the estimated intercept than variability between the period of sample collection. Due to the large number of women with only one measurement here, the within-woman correlation coefficient is unlikely to result in an accurate estimate of the correlation of HIV RNA measurements within a woman over pregnancy; this will be estimated in the following section in the sensitivity analysis involving women with at least two measurements.

3.3.5 Sensitivity analysis

The majority of women (65%) in this analysis had only one measurement in pregnancy, limiting any inferences which can be made on the shape of HIV RNA curves in pregnancy and on the accuracy of the within-woman correlation estimate. Additionally, 79% (173/218) of women with only one measurement in pregnancy had their measurement taken at delivery; the resulting estimate of the slope of change in HIV RNA may therefore be confounded by factors influencing when in pregnancy this measurement was made.

Therefore a sensitivity analysis including only the 80 women with at least two HIV RNA measurements in pregnancy and with available information on all variables was carried out. The model included the same covariate and interaction terms but did not include a

random effect term for the sampling period as the estimate for this term was extremely small (σ<0.05) and its inclusion led to a variance-covariance matrix for the random effects which was not positive-definite. The mean HIV RNA at delivery was estimated to be 3.86 $\log_{10}$ copies/ml (95%CI 3.44-4.28), higher than in the model which included all women, while the magnitude of the estimated change in HIV RNA over pregnancy was smaller, at -0.013 $\log_{10}$ copies/ml per week (95%CI -0.025, -0.0004). The largest difference in the parameter estimates from the model which included all women was that of the race effect; black women had estimated viral loads of 0.11 $\log_{10}$ copies/ml (-0.39, 0.60) above that of white women at delivery, but this was not significantly different from zero. Similar to the model presented in Table 3.4, the slope of change in viral load over pregnancy for black women was positive and significantly different to that for white women (0.030 $\log_{10}$ copies/ml [95%CI 0.0010-0.055]), but the change in viral load over pregnancy for black women was not significantly different to zero (0.028 $\log_{10}$ copies/ml [95%CI -0.45, 0.51]). The between- and within- women standard deviations were estimated to be 0.58 and 0.49, respectively. The within-woman correlation can be calculated as $\sigma^2_{\text{intercept}} / \left[\sigma^2_{\text{intercept}} + \sigma^2_{\text{residual}} \right]$: therefore, the estimated correlation of $\log_{10}$ HIV RNA within a woman over the 24 weeks of pregnancy observed was 0.66 (95%CI 0.50-0.79).

As mentioned in section 3.2.2, a nested mixed effects model was used to study HIV RNA changes in pregnancy, with midpoints imputed for measurements recorded at the assay lower limit of quantification. The LME model which takes into account the left-censoring using a conditional likelihood approach has been shown to be more accurate for analysing changes in left-censored HIV RNA measurements (Jacqmin-Gadda *et al.* 2000) (see also Section 5.2.2). Therefore a sensitivity analysis was carried out using this

method instead, with the sampling period included in the model as a covariate. Although the resulting estimate for mean HIV RNA at delivery was higher (4.32 $\log_{10}$ copies/ml [95%CI 3.63-4.99]), the estimate for change in HIV RNA was similar (-0.022 $\log_{10}$ copies/ml per week [95%CI -0.037, -0.007]). The conditional likelihood model resulted in a relatively more marked impact on the intercept estimate which is most likely due to the large number of undetectable measurements at delivery.

3.4 Key points

- In this subgroup of untreated pregnant women, levels of HIV RNA viral load did not remain constant over pregnancy.
- HIV RNA viral load over pregnancy was estimated to decrease for white women and to remain generally stable for black women.
- Black women had significantly higher viral loads at delivery than white women
- Levels of HIV RNA in pregnancy were associated with baseline maternal CD4 cell counts and on the type of blood sample used
- There was considerable variability between samples collected before and during the period of routine viral load testing, justifying the use of a nested random-effects model.
- Accounting for left-censoring of HIV RNA levels with midpoints led to heteroscedasticity of the within-group errors. However, the addition of a variance function to model the within-group variability of undetectable values significantly improved the fit of the model.
- In a model including women with only two measurements, HIV RNA viral load over pregnancy was also estimated to decrease for white women, but the magnitude of decrease was smaller.
- The estimated correlation of $\log_{10}$ HIV RNA within a woman over the second and third trimesters of pregnancy was estimated to be 0.66.

Chapter 4 Time to undetectable HIV RNA viral load after initiation of HAART

4.1 Introduction

Plasma HIV RNA level is the pre-eminent risk factor for MTCT of HIV (Mofenson *et al.* 1999;European Collaborative Study 2005b). HAART in resource-rich settings has substantially reduced MTCT rates through successful suppression of HIV RNA in pregnant HIV-infected women (European Collaborative Study 2005b;Cooper *et al.* 2002). While an increasing proportion of HIV-infected pregnant women in such settings are identified and treated before pregnancy, a substantial minority are diagnosed antenatally and start ARV drugs to delay disease progression and/or prevent MTCT for the first time in pregnancy (European Collaborative Study 2005b). In many Western European countries, these women are increasingly likely to have acquired HIV heterosexually and to be from countries with generalised epidemics, mostly sub-Saharan Africa (European Collaborative Study 2006). No clinical trials in resource-rich settings have addressed the question of which regimens are more effective for optimal viral response in ARV-naïve pregnant women.

This analysis was carried out to determine whether choice of initial HAART regimen in HIV-infected pregnant women is associated with the time to achieving undetectable viral load by delivery.

4.2 Methods specific to this analysis

4.2.1 Data

The analysis was restricted to the 240 women who were ART-naïve at conception, 153 women identified for the first time in pregnancy and 87 before pregnancy. Of the 87 women diagnosed before pregnancy, 46 were known not to have previously received any PMTCT prophylaxis (either because of no prior pregnancies or documented non-receipt of ARVs in previous ECS pregnancies) and 41 were not on ARVs at conception with no documented prior ARV use. Viral load at the time of initiating HAART has been shown to be associated with the probability of reaching a viral load < 500 copies/ml, with higher pre-HAART viral loads less likely to achieve this level (Powderly *et al.* 1999;Mocroft *et al.* 1998). Other studies examining the time to undetectable HIV RNA after initiation of different HAART regimens included patients with an HIV RNA measurement within 6 months before initiation (Phillips *et al.* 2001;Ghani *et al.* 2001). However, in light of the findings in Chapter 3 of non-stable HIV RNA levels in pregnancy and given that this study was limited to pregnancy, the eligibility criterion was confined to women with a HIV RNA measurement within 6 weeks before initiation and at least one subsequent measurement thereafter.

There were 759 HIV RNA measurements available on the 240 women included in the analysis. Classification of undetectable viral load (viral suppression) was based on the assay lower limit of quantification. Of the 759 measurements, 561 (74%) were measured using an ultrasensitive assay (with a quantification limit ≤50 copies/ml). HAART was defined as a regimen of three or more antiretroviral drugs consisting of a NRTI backbone and including NVP – a NNRTI, or a PI.

After adjusting for baseline viral load, race (non-black vs. black), HAART type (PI-containing vs. NVP-containing), baseline CD4 count (<200, 200-499 or ≥500 cells/mm^3), trimester at initiation (1st trimester, 2nd trimester or 3rd trimester), timing of HIV diagnosis (pre-pregnancy vs. antenatal), history of IDU (Y/N), year of delivery (1998-2000, 2001-2002 or 2003-2004) were considered in the Weibull model and propensity score model, respectively. These variables were deemed important potential confounders as they are known to or have the potential to be related to the time taken to achieve undetectable HIV RNA levels (Ghani *et al.* 2001;Mocroft *et al.* 2006a;Matthews *et al.* 2002;Phillips *et al.* 2001).

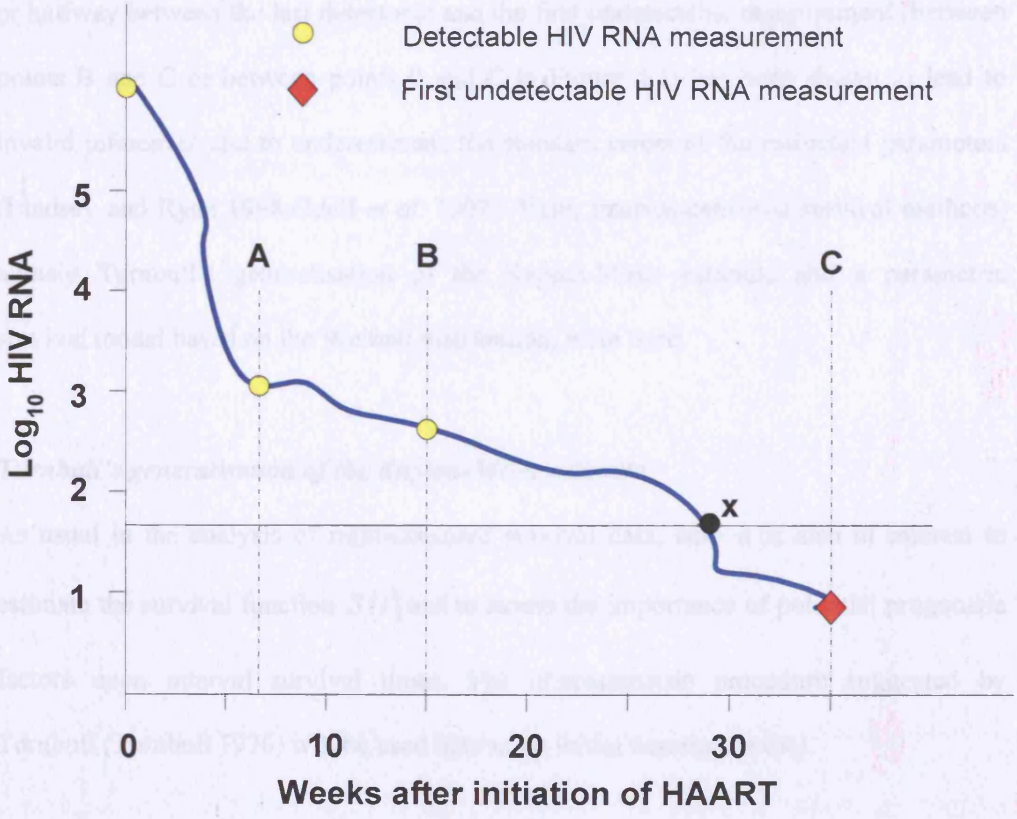
4.2.2 Statistical methods

In observational studies, the timing of HIV RNA measurements after treatment initiation occurs at various times and with various frequencies. Thus, a common approach to compare different HAART regimens is to examine the times to events using survival analysis methods rather than comparing absolute changes in levels of virological and immunological markers. A frequently used measure of success of treatment is the time taken from initiation of treatment until HIV RNA reaches undetectable levels (Matthews *et al.* 2002;Ghani *et al.* 2001;Phillips *et al.* 2001). In this analysis, the association between initial HAART regimen and maternal baseline characteristics with viral suppression was examined by comparing the times taken from treatment initiation to the first undetectable HIV RNA measurement, through to the time of delivery.

The analysis of time to undetectable HIV RNA after initiation of HAART here results in a complicated censoring pattern. Figure 4.1 shows simulated HIV RNA data for patients

after initiating HAART, where the time point 0 corresponds to the time of initiation. If a subject's true endpoint is known only to occur in the interval [0, B), i.e. it has not occurred at the time of the last measurement, then the resulting survival time is right censored. Where a subject's true endpoint, marked as **x** in figure 4.1, is known only to occur in the interval [B, C], the resulting survival time is interval-censored, that is **x** is known only to have occurred between B and C. Finally, the survival time is left-censored when **x** is known only to have occurred in the interval [0, C] and is an exact time when the endpoint is known to occur at **x.**

Figure 4.1 Figure to illustrate censoring pattern from analysing time to undetectable HIV RNA after initiation of HAART



Therefore for women not achieving undetectable levels, the time was right-censored at the last measurement at or before delivery; for those reaching undetectable virus, the time was left-censored if the endpoint was known only to have occurred between initiation and the first measurement, and interval-censored if the endpoint occurred between any two measurements following initiation. There were no exact survival times as the true date of the first undetectable measurement was unknown for all women.

Standard time-to-event methods such as the Product limit estimator developed by Kaplan and Meier and the Cox proportional hazards models can only accommodate right-censoring; the use of these methods with interval-censored data, assuming that the true time **x** of the event occurs either at the time of the first undetectable measurement or halfway between the last detectable and the first undetectable measurement (between points B and C or between points 0 and C in Figure 4.1) has been shown to lead to invalid inferences and to underestimate the standard errors of the estimated parameters (Lindsey and Ryan 1998;Odell *et al.* 1992). Here, interval-censored survival methods, namely Turnbull's generalisation of the Kaplan-Meier estimate and a parametric survival model based on the Weibull distribution, were used.

Turnbull's generalisation of the Kaplan-Meier estimate

As usual in the analysis of right-censored survival data, here it is also of interest to estimate the survival function $S(t)$ and to assess the importance of potential prognostic factors upon interval survival times. The nonparametric procedure suggested by Turnbull (Turnbull 1976) will be used here as an initial descriptive tool.

The modified Product-limit estimator suggested by Turnbull (Turnbull 1976) has no closed form and is based on the following iterative procedure (Klein and Moeschberger 1997). Let $T_i\left(i=1,2,...,n\right)$ be the time until event which is known only to occur between an interval $(L_i, R_i]$ where $L_i < T_i \leq R_i$. Note that for right-censored observations $R_i = \infty$ and for left-censored observations $L_i = 0$. Let $0=t_0<t_1<t_2<...<t_m$ be a partition which includes all the points L_i and R_i for $i = 1,...,n$. For the ith observation, a weight α_{ij} is calculated, such that $\alpha_{ij}=1$ if the interval $\left(t_{j-1},t_j\right]$ is contained in the interval $(L_i, R_i]$, and 0 otherwise. α_{ij} is an indicator of whether the event which occurs in the interval $(L_i, R_i]$ could have occurred at t_j. Turnbull's algorithm is as follows:

Step 0: Produce an initial Kaplan-Meier estimator of the survival function $S\left(t_j\right)$ at each t_j, by taking the end points of intervals R_i as the survival times.

Step 1: Compute the probability of an event occurring at time t_j as

$$p_j = S\left(t_{j-1}\right) - S\left(t_j\right) \quad j=1,...,m$$

Step 2: Estimate the number of events occurring at t_i by

$$d_i = \sum_{i=1}^{n} \frac{\alpha_{ij} p_j}{\sum_{k=1}^{m} \alpha_{ij} p_k} \qquad j=1,...,m.$$

Step 3: Compute the estimated number of subjects at risk at time t_i as

$$Y_i = \sum_{k=j}^{m} d_k .$$

Step 4: Compute the updated Kaplan-Meier estimator based on the estimated number of deaths and number at risk found in steps 2 and 3. If the updated estimate of $S\left(t_j\right)$ is

close to the old version for S for all t_i's then stop the iterative process, otherwise repeat steps 1-3 using the updated estimate of S.

At present, this algorithm is not available as a standard function in the R statistical software package, and an algorithm developed by Giolo was used to obtain stratified survival curves and associated survival probabilities (Giolo 2004). The R code for this is given in the Appendix D. 95% CIs for the survival probabilities were calculated using the adjusted bootstrap percentile method with 1000 replications (Davison and Hinkley 1997).

Parametric regression models for survival data

The standard Cox proportional hazards model does not require any specific distributional assumptions about the shape of the survival function. In fact a parametric (linear) form is assumed only for the covariate effects, the baseline hazard is unspecified and is not actually estimated using the partial likelihood approach (Klein and Moeschberger 1997). Although this is one of the advantages of this type of model over parametric regression models, the resulting estimates are not as efficient as maximum likelihood estimates for correctly specified parametric survival models. Additionally, as a full likelihood is not specified, the standard Cox model can not allow for left or interval-censored data (Klein and Moeschberger 1997). It has also been shown elsewhere that when data are heavily censored, making assumptions about when events occurred (e.g. assuming midpoints) and using techniques such as Cox regression can lead to inaccurate conclusions (Lindsey and Ryan 1998). It has been suggested that parametric models are an appropriate alternative method, which can be highly satisfactory in their performance (Lindsey and Ryan 1998). Parametric regression

models by definition assume a specific distribution for the shape of the survival function and are able to incorporate left-, right- and interval-censored data. Therefore, due to the nature and amount of censoring in this study (no women had exactly observed event times and a large majority had left- and interval-censored event times), a parametric regression model was preferred over the Cox model.

Parametric survival models can usually be represented as an accelerated-failure time (AFT) model or a linear model in log time. Let X denote the time to event and $\mathbf{Z}$ a vector of fixed-time covariates. The AFT model states that the survival function for an individual with covariate $\mathbf{Z}$ at time x can be is defined by

$$S(x \mid \mathbf{Z}) = S_0\left[\exp\left(\boldsymbol{\theta}^T\mathbf{Z}\right)x\right], \text{ for all } x, \quad (4.1)$$

where $\boldsymbol{\theta}^T = (\theta_1,...,\theta_p)$, where $j = 1,...,p$ is a vector of regression coefficients and S_0 is the baseline survival function. The factor $\exp\left(\boldsymbol{\theta}^T\mathbf{Z}\right)$ is the acceleration factor and describes how a change in the covariate value changes the time scale from the baseline time scale, i.e. whether the effect of changes in the covariates results in a shrinking or stretching of the time to event. An implication of this model is that the hazard rate for an individual with covariate $\mathbf{Z}$ is related to the baseline by

$$h(x \mid \mathbf{Z}) = \exp\left(\boldsymbol{\theta}^T\mathbf{Z}\right)h_0\left[\exp\left(\boldsymbol{\theta}^T\mathbf{Z}\right)x\right], \text{ for all } x.$$

The second representation is the log-linear form between log-time and the covariate values, such that

$$Y = \ln X = \mu + \boldsymbol{\gamma}^T\mathbf{Z} + \sigma W,$$

where $\gamma^T = (\gamma_1, ..., \gamma_p)$ is a vector of regression coefficients and W is the random error distribution. By taking S_0 to be the survivor function of the random variable $\exp(\mu + \sigma W)$, the linear log-time model is equivalent to the AFT model with $\theta = -\gamma$.

Various parametric distributions can be used for the random term W or S_0, including the log-logistic and the exponential distributions. Here a Weibull model is preferred as it is flexible, allowing for monotone increasing, decreasing or constant hazards and has been shown to be more appropriate than the Cox model in the analysis of data with a large proportion of left- and interval-censored observations (Lindsey and Ryan 1998). It is also the only parametric survival model which has both a proportional hazards and an AFT representation, facilitating the interpretation of the estimated model coefficients (Lindsey and Ryan 1998;Klein and Moeschberger 1997). Additionally the assumption of proportional hazards does not need to hold in a Weibull model and relative survival times can be obtained from the fitted model, additional to relative hazards (Cox *et al.* 2007).

The baseline survival and hazard function for the Weibull distribution can be represented as

$$S_0(x) = \exp(-\lambda x^{\alpha})$$

$$h_0(x) = \lambda \alpha x^{\alpha - 1},$$

respectively. Note that if α=1, the Weibull distribution reduces to an exponential distribution.

The linear log-time representation of the Weibull model is then

$$Y = \ln X = \mu + \gamma^T \mathbf{Z} + \sigma W,$$

where W has the standard extreme value distribution (the error distribution for the Weibull model). This model also leads to the proportional hazards model for X with the baseline hazard function described above and can be expressed as

$$h(x \mid \mathbf{Z}) = \left(\alpha \lambda x^{\alpha-1}\right) \exp\left(\boldsymbol{\beta}^T \mathbf{Z}\right).$$

This representation is linked to the coefficients in the linear log-time model as $\alpha = 1/\sigma$, $\lambda = \exp\left(-\mu/\sigma\right)$ and $\beta_j = -\gamma_j/\sigma$. The AFT representation of this model is the same as in equation 4.1.

Likelihood construction for interval-censored data

Assuming a full-likelihood approach, the likelihood function for survival data can be constructed to account for right, left and interval-censoring through consideration of the information each observation provides (Klein and Moeschberger 1997).

For a specific individual under study, assume that X denotes the time to event, C_r and C_l denote a right-censored and left-censored time, respectively, and the interval $(L_i, R_i]$ denotes the left and right endpoint of an interval-censored time. Further, the X's are assumed to be independent and identically distributed with probability density function $f(x)$ and survival function $S(x)$. An observation corresponding to an exact survival time then provides information on the probability that the event occurs at that time, which is approximately equal to $f(x)$, the density function of X at this time. For a right-censored observation, it is only known that the event time is larger than this time, so the information is equivalent to $S(C_r)$, the survival function evaluated at that study

time. For a left-censored observation, it is known that the event has already occurred, so the contribution to the likelihood is the cumulative distribution function, $1-S(C_l)$, evaluated at that study time. Finally, for interval-censored data it is known that the event occurred within an interval so the information is $\left[S(L_i)-S(R_i)\right]$, the probability that the event time is in this interval.

The likelihood function for exact and right-, left- and interval-censored survival times can then be constructed by putting together the above components, such that

$$L \propto \prod_{i \in E} f(x_i) \prod_{i \in R} S(C_r) \prod_{i \in L} (1-S(C_l)) \prod_{i \in I} \left[S(L_i)-S(R_i)\right],$$

where E is the set of event times, R is the set of right-censored observations, L is the set of left-censored observations and I the set of interval-censored observations.

Propensity scores

Randomised controlled trials are recognised as the gold standard for making treatment comparisons, as randomisation of patients to each treatment group means that on average, each patient has a similar prognosis at the start of the study and treatment effects can be singled out. In observational studies however, there is no control over treatment assignment and therefore large differences on observed covariates in the two groups may exist at baseline, which could lead to biased treatment effects. Traditional methods of adjustment (matching, stratification, adjustment in regression models) remove some of this bias but can use only a limited number of covariates for adjustment. The propensity score can instead be used to reduce this bias without this limitation, as it provides a scalar summary score of all of the covariate information (D'Agostino, Jr. 1998).

The propensity score for an individual is defined as the probability of being treated with one treatment over another (e.g. treatment vs. control, PI-regimen vs. NNRTI regimen), conditional on the individual's covariates at baseline. It can be used to reduce bias in treatment comparisons with observational data by balancing the covariates in the two groups.

Rosenbaum and Rubin (Rosenbaum and Rubin 1983), define the propensity score for subject i (i = 1, ..., N) as the conditional probability of assignment to a particular treatment ($Z_i = 1$) versus control ($Z_i = 0$) , given a vector of observed covariates, x_i:

$$e(x_i) = pr(Z_i = 1 \mid X_i = x_i)$$

where it is assumed that given the X's, the Z_i are independent:

$$pr(Z_1 = z_1, ..., Z_N = z_N \mid X_1 = x_1, ..., X_N = x_N) = \prod_{i=1}^{N} e(x_i)^{z_i} \{1 - e(x_i)\}^{1-z_i} .$$

As a crude function of the covariates, the propensity score can be perceived as a balancing score, *b(X),* defined as "a function of the observed covariates X such that the conditional distribution of X given *b(X)* is the same for treated *(Z=1)* and control *(Z=0)* individuals (D'Agostino, Jr. 1998). Therefore, matching, stratification or regression adjustment on the propensity score will produce unbiased estimates of treatment effects, if the treatment assignment is ignorable, given the observed covariates. However, unmeasured variables can still be unbalanced between the two treatment groups even

after incorporating the propensity score (Austin *et al.* 2005); in this study this includes among other unmeasured variables, data on adherence and time of seroconversion.

When complete data are available, the propensity score can be estimated using discriminant analysis or logistic regression. Although both techniques will lead to estimates of probabilities of treatment assignment given a set of observed covariates, discriminant analysis necessitates the added assumption of the covariates having a multivariate normal distribution (conditional on Z).

The propensity scores were used here for regression adjustment, i.e. using the probability (i.e. the propensity score) that any given patient is assigned one treatment over another, to adjust regression estimates of treatment effects. The estimation and use of the score in this analysis follows the recommendations from D'Agostino (D'Agostino, Jr. 1998) and is summarised below.

1) With a binary treatment outcome, estimate the propensity score using logistic regression, with covariates entered into the model and retained if significant at $p<0.5$ in a stepwise selection procedure.
2) Obtain estimated propensity scores from the fitted model.
3) Determine strata cut-offs using quintiles of the estimated propensity scores for the combined group.
4) Incorporate this 5-level factor into the model of interest, regardless of its significance, together with other confounders whose association with the treatment outcome is to estimated.

Covariance adjustment for all variables which appear in the propensity score model has been shown to result in a similar treatment effect (D'Agostino, Jr. 1998). However, an advantage of this approach over adjustment for all variables is that although a large set of background covariates can be used to estimate the propensity scores, only a subset of these covariates and the propensity score itself are needed in the regression adjustment. As the goal of the propensity score model is to obtain the most accurate probability of treatment assignment, covariates in this model can be included although they do not appear in the Weibull regression model. Although, by definition a variable will only be a confounder if it is associated with both treatment assignment and outcome, Austin and colleagues found that covariance adjustment for variables which were true confounders only resulted in biased estimation of the treatment effect hazard ratio, compared to models which adjusted for all variables or those which adjusted for variables associated with the event outcome only (Austin *et al.* 2007). Therefore, in order to reduce the possibility of a biased treatment effect estimate, to attempt to balance treatment assignment by observed covariates and to create a "quasi-randomised" trial setting, the final model was adjusted for both the propensity score and a subset of these covariates.

4.2.3 Model selection procedure

The model selection procedure follows that described in section 3.2 but was based on assessment of the deviance statistic for the logistic regression model and the LRT for the survival model. However, baseline viral load, stratified into tertiles, was included in the model a priori as the outcome of interest is expected to be linked to baseline viral load; therefore the model selection procedure involved the examination of improvements from the inclusion of additional covariates from a model including viral load only. Baseline CD4 cell counts were included in the final model regardless of the

statistical significance of the improvement in model fit as is it also expected to be related to the outcome of interest and is an indicator of HIV disease status. The propensity of being treated with NVP-based HAART over PI-based HAART, estimated through a logistic regression model was then included in the final adjusted models after stratification into quintiles (D'Agostino, Jr. 1998).

The Weibull regression model was carried out using the *intcens* function in Stata version 9.1 (StataCorp LP, College Station, Texas) and the *survreg* package in R version 2.2.0 (R Development Core Team 2006). The logistic regression model was fitted using R.

4.3 Results

Two hundred and forty pregnant HIV-infected women met the inclusion criteria, out of 1,346 women on antenatal HAART delivering between 1997 and 2004. The characteristics of these women are given in Table 4.1. Most women (59%) were black, of whom 90% were born in sub-Saharan Africa (Table 4.1).

PI-based HAART was initiated in 156 (65%) women, with the remaining 84 (35%) receiving an NNRTI-based regimen (all NVP). Most women received a ZDV/3TC NRTI backbone (Table 4.1). Eighty percent (n =125) of PI-based regimens included NFV, with the remaining containing lopinavir/ritonavir (n=4), ritonavir (n=13), indinavir (n=8) or saquinavir (n=6). The proportion of women receiving NVP-based HAART increased from 25% (16/64) in 1997-2000 to 35% (37/105) and 44% (31/71) in 2001-2002 and 2003-2004, respectively (p=0.02). Among the NVP group, 61 (73%) women had CD4 counts >250 cells/mm^3 at initiation, of whom 57 (93%) delivered

Table 4.1 Characteristics of the 240 treatment-naïve study women at initiation of HAART and number of women reaching endpoint of undetectable HIV RNA viral load by delivery

Characteristic	Value (Column %)	Number of women reaching undetectable viral load (Row %)
Race		
Non-black	96 (41)	63 (66)
Black	141 (59)	110 (78)
Unknown	3	2
Region of birth*		
Europe	75 (32)	46 (61)
The Americas	23 (10)	18 (78)
Asia	8 (3)	7 (88)
Northern Africa	6 (2)	5 (83)
Eastern Africa	44 (19)	33 (75)
Southern Africa	2 (1)	1 (50)
Central Africa	40 (17)	31 (78)
Western Africa	39 (16)	32 (82)
Unknown	3	2
Age at delivery (years)		
Median (IQR)	29 (25-33)	
15-19	10 (4)	6 (60)
20-29	113 (48)	86 (76)
30-39	107 (45)	74 (69)
≥40	8 (3)	7 (88)
Unknown	2	2

* - regions of Africa were defined according to the United Nations groupings

Table 4.1 contd.

Characteristic	Value	Number of women reaching undetectable viral load
	(Column %)	(Row %)
History of IDU		
Non-IDU	215 (91)	159 (74)
IDU	21 (9)	12 (57)
Unknown	4	4
Timing of HIV diagnosis		
Antenatal	153 (64)	113 (74)
Pre-pregnancy	87 (36)	62 (71)
HIV RNA (log_{10} copies/ml)		
Median (IQR)	4.16 (3.62-4.58)	
≥5	20 (8)	10 (50)
4-4.99	125 (52)	88 (70)
3-3.99	74 (31)	58 (78)
<3	21 (9)	19 (90)
CD4 count (cells/mm^3)		
Median (IQR)	328 (210– 480)	
<200	48 (22)	29 (60)
200-499	124 (56)	95 (77)
≥500	48 (22)	38 (79)
Unknown	20	13
HAART category		
PI-based	156 (65)	111 (71)
NVP-based	84 (35)	64 (76)
NRTI backbone		
zidovudine/lamivudine	211 (88)	160 (76)
Other dual combination	29 (12)	15 (52)

Table 4.1 contd.

Characteristic	Value (Column %)	Number of women reaching undetectable viral load (Row %)
Stage of pregnancy at initiation (weeks' gestation)		
Median (IQR)	23 (18–27)	
First trimester	14 (6)	12 (86)
Second trimester	168 (70)	129 (77)
Third trimester	58 (24)	34 (59)
Gestational age at delivery		
Median (IQR)	38 (37-39)	
Delivery at term (≥37 weeks)	187 (78)	148 (79)
Premature (<37 weeks)	53 (22)	27 (51)

before February 2004 (when NVP prescribing information changed (Public Health Service Taskforce 2006b)).

Characteristics by race

There were no differences between black and non-black women with respect to type of HAART (34% [48/141] vs. 35% [34/96] received NVP-based HAART respectively; p=0.94), the distribution of baseline median HIV RNA viral load (4.13 $\log_{10}$ copies/ml [IQR 3.54-4.60] vs. 4.20 $\log_{10}$ copies/ml [IQR 3.76-4.56]; p=0.18) and baseline median CD4 cell counts (313 cells/mm^3 [IQR 210-449] vs. 370 cells/mm^3 [IQR 231-528]; p=0.12). However, black women tended to start treatment later than non-black women (median 24 weeks' gestation [IQR 20-28] vs. median 20.5 weeks [IQR 15-27]; p=0.007).

Figure 4.2 shows the distribution of HIV RNA measurements at initiation by gestational weeks, together with baseline median viral loads and CD4 counts. Although numbers of women were too small to explore HIV RNA measurements by gestational week, time periods were categorised into periods of five weeks to better utilise the information without grouping measurements into trimesters of pregnancy. Twenty-four percent (39/165) of women starting HAART in the first or second trimester had CD4 counts below 200 cells/mm^3 compared to 16% (9/55) of those starting in the third trimester (p=0.35).

Characteristics by HAART category

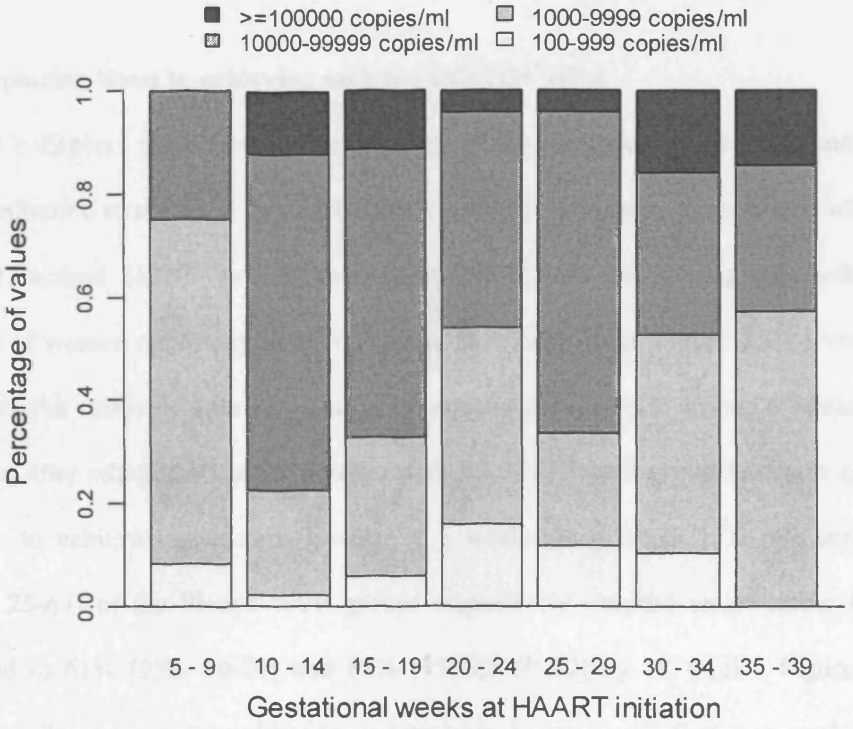
Table 4.2 shows the characteristics of women by HAART category. Overall, from initiation of treatment, the median number of virological measurements per woman was three (range 2-7), with a similar interval between successive tests for the two treatment groups (Table 4.2). Although timing of HAART initiation and baseline HIV RNA were similar between treatment groups, the PI-based group had significantly lower baseline CD4 counts (Table 4.2). This suggests the possible presence of confounding by indication with women with lower baseline CD4 counts preferentially initiated on PI-based HAART rather than a NVP-based regimen; adjusting for baseline CD4 count category and for treatment assignment with the propensity score will reduce some of this bias. The proportion of women delivering prematurely was similar between the treatment groups, however women receiving NVP-based HAART delivered slightly earlier than women receiving PI-based HAART (Table 4.2); this was chiefly due to the larger proportion of elective caesarean section deliveries being carried out at or before 37 weeks' gestation in women who received NVP-based HAART (74% [57/77] for NVP-based vs. 57% [84/148] for PI-based HAART; χ^2=5.74, p=0.02).

Table 4.2 Characteristics by HAART category

Characteristic	PI-based HAART (n =156)	NVP-based HAART (n =84)	p value†
NRTI backbone			
zidovudine/lamivudine	139 (89%)	72 (86%)	
Other dual combination	17 (11%)	12 (14%)	0.58
Median time of initiation (gestational weeks)	23 (IQR 18-27)	21.5 (IQR 16-28)	0.57
Median baseline HIV RNA viral load (log_{10} copies/ml)	4.18 (IQR 3.60-4.58)	4.08 (IQR 3.71-4.54)	0.58
Median baseline CD4 count (cells/mm^3)	305 (IQR 190-452)	355 (IQR 277-506)	0.02
Median number of measurements	3 (IQR 2-3)	3 (IQR 2-3)	0.77
Median interval between successive HIV RNA tests (weeks)	7.5 (IQR 4-10)	6 (IQR 4-10)	0.07
History of IDU			
Non-IDU	140 (92%)	75 (90%)	
IDU	13 (8%)	8 (10%)	0.96
Timing of HIV diagnosis			
Antenatal	103 (66%)	50 (60%)	
Pre-pregnancy	53 (34%)	34 (40%)	0.37
Median gestational age at delivery (weeks)	38 (range 25-42)	37 (range 23-41)	0.003
Premature delivery (<37 weeks)			
No	126	61	
Yes	30	23	0.20

† - p values were calculated with the Chi-squared or Mann-Whitney test as appropriate

Figure 4.2 Distribution of baseline viral load measurements and median baseline viral load and CD4 cell count at time of HAART initiation in pregnancy



	5 - 9	10 - 14	15 - 19	20 - 24	25 - 29	30 - 34	35 - 39
n	12	49	52	50	51	19	7
Median VL(log_{10} copies/ml)	3.92	4.50	4.19	3.76	4.17	3.85	3.82
(IQR)	(3.40-4.09)	(4.03-4.75)	(3.85-4.71)	(3.22-4.32)	(3.82-4.48)	(3.25-4.67)	(3.75-4.28)
Median CD4 (cells/mm^3)	392	321	332	313	322	450	500
(IQR)	(271-512)	(171-441)	(209-477)	(210-475)	(247-390)	(197-557)	(325-600)

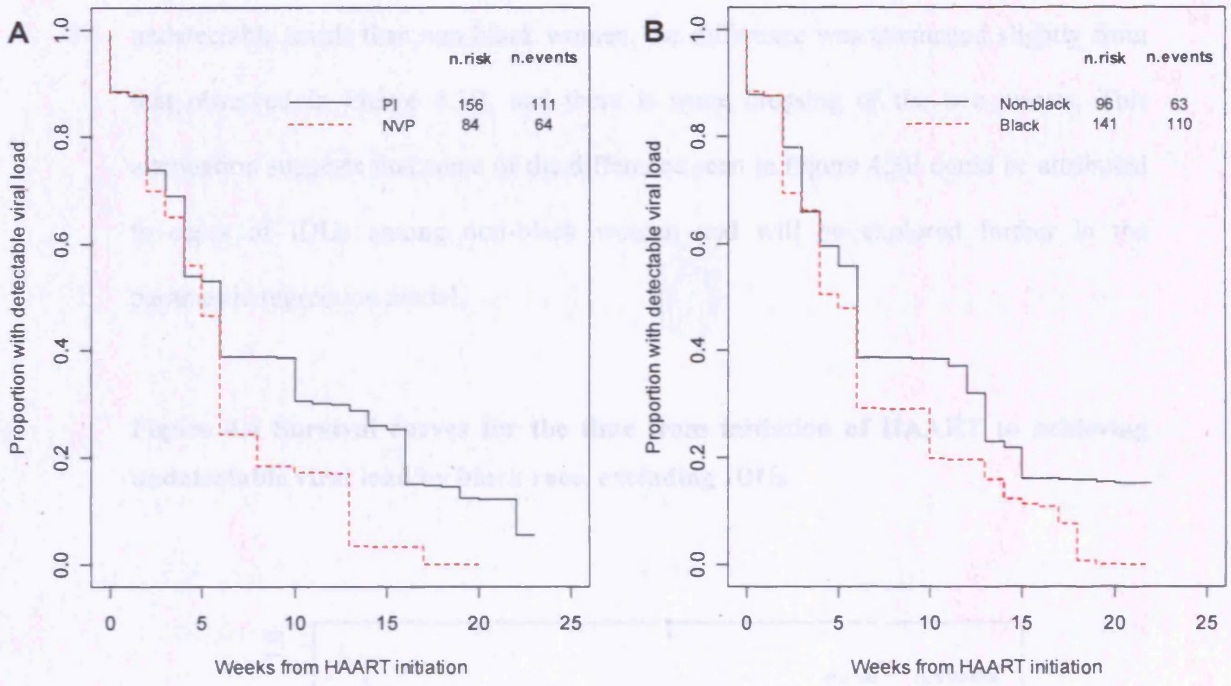
The median gestational age at delivery was 38 weeks (range 23-42) and 73% (175/240) of women had undetectable viral loads by this time; Table 4.1 shows the number reaching this outcome by maternal characteristics. The proportion of women achieving undetectable viral load did not differ significantly by treatment group (71% [111/156] for PI-based and 76% [64/84] for NVP-based HAART; p=0.49). It should be noted here that as most women reached viral suppression, the relative hazards (RH) obtained from subsequent regression models will reflect the rapidity with (rate at) which this occurs,

rather than solely the probability of the event occurring; larger RHs will therefore indicate more rapid attainment of viral suppression.

4.3.1 Exploring times to achieving undetectable HIV RNA

Figure 4.3 displays the proportions of women achieving undetectable viral load from time of initiation stratified by HAART category and by black race, as estimated with the Turnbull method (Klein and Moeschberger 1997;Giolo 2004), together with the numbers of women remaining in the risk set at each time point. Figure 4.3A reveals an overall similar survival curve between both treatment groups up to around 6 weeks after initiation, after which the curves diverge with the NVP-based group having a greater tendency to achieve undetectable levels; at 5 weeks 46% [95%CI 33-60] and 44% [95%CI 25-67] of the PI and NVP groups respectively reached undetectable levels, compared to 61% [95% 50-71] and 82% [95%CI 69-92] by 10 weeks. Figure 4.3B reveals similar proportions achieving undetectable levels in the first few weeks after initiation among non-black and black women, after which the black women were observed to have a greater tendency to achieve undetectable levels.

Figure 4.3 Survival curves for the time from initiation of HAART to achieving undetectable viral load by: (A) treatment category; (B) black race



	n at risk (Fig 4.3A)			*n* at risk (Fig. 4.3B)	
Weeks from initiation	***PI***	***NVP***		***Non-black***	***Black***
0	156	84		96	141
5	137	73		81	126
10	96	46		58	81
15	55	19		32	40
20	25	6		15	13
25	4	2		4	2

Figure 4.4 reveals the survival curves for black and non-black women excluding IDUs; the survival curves were similar in the first few weeks after initiation, and although there was some evidence of black women having a greater tendency to achieve undetectable levels than non-black women, the difference was attenuated slightly from that observed in Figure 4.3B, and there is some crossing of the two curves. This attenuation suggests that some of the difference seen in Figure 4.3B could be attributed to cases of IDUs among non-black women and will be explored further in the parametric regression model.

Figure 4.4 Survival curves for the time from initiation of HAART to achieving undetectable viral load by black race, excluding IDUs

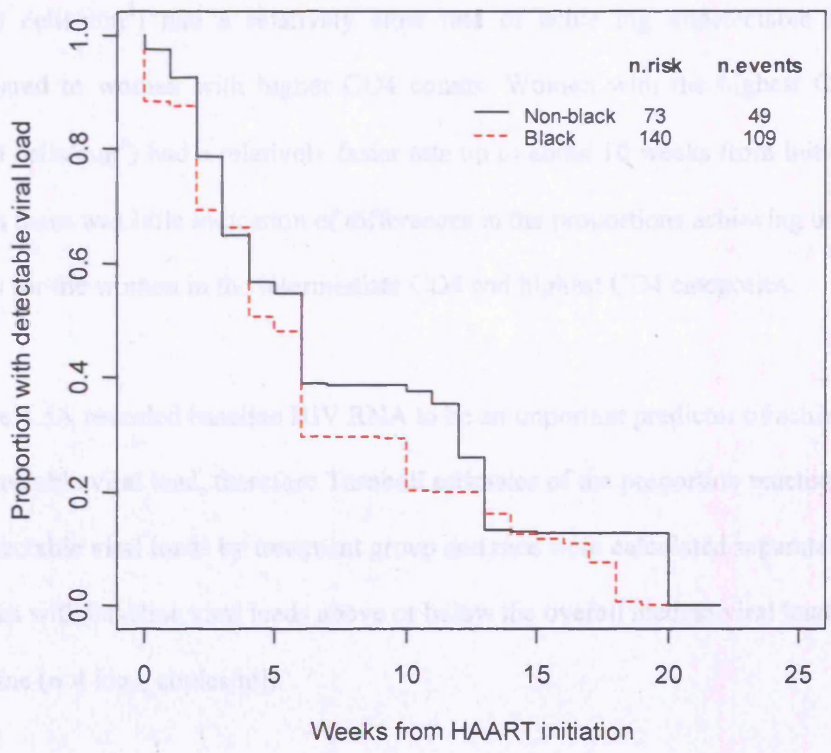
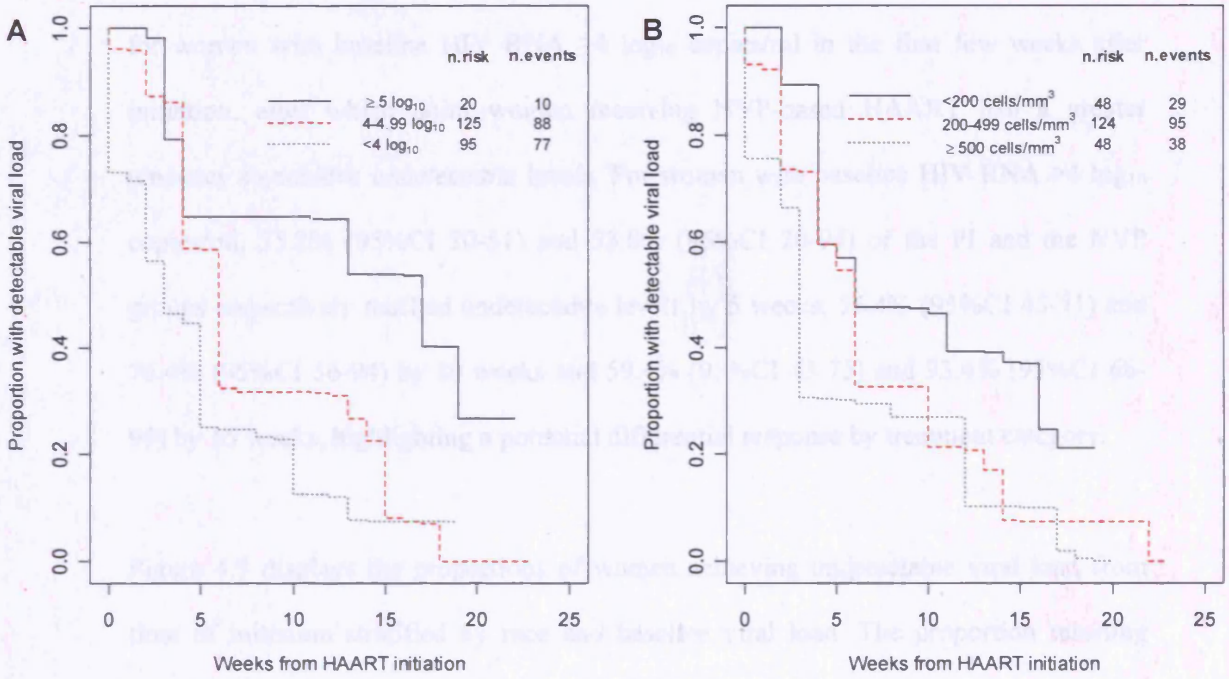


Figure 4.5A shows the Turnbull estimates of the proportion achieving undetectable viral load according to baseline HIV RNA levels, and revealed marked differences in the rates of achieving undetectable viral load between these categories. Overall women with the highest baseline HIV RNA (≥ 5 $\log_{10}$ copies/ml) had a lesser tendency to achieve undetectable levels than those with lower baseline levels, although the curves for women in this group were similar to those with baseline HIV RNA 4-4.99 $\log_{10}$ copies/ml for the first few weeks after initiation. Women with the lowest baseline HIV RNA (<4 $\log_{10}$ copies/ml) had the fastest rate of achieving undetectable HIV RNA levels.

Figure 4.5B shows the Turnbull estimates of the proportion achieving undetectable viral load according to baseline CD4 cell counts. Women with low baseline CD4 counts (<200 cells/mm^3) had a relatively slow rate of achieving undetectable HIV RNA compared to women with higher CD4 counts. Women with the highest CD4 counts (≥ 500 cells/mm^3) had a relatively faster rate up to about 10 weeks from initiation, after which there was little indication of differences in the proportions achieving undetectable levels for the women in the intermediate CD4 and highest CD4 categories.

Figure 4.5A revealed baseline HIV RNA to be an important predictor of achieving an undetectable viral load, therefore Turnbull estimates of the proportion reaching undetectable viral loads by treatment group and race were calculated separately for women with baseline viral loads above or below the overall median viral load at baseline (≈ 4 $\log_{10}$ copies/ml).

Figure 4.5 Survival curves for the time from initiation of HAART to achieving undetectable viral load by (A) baseline viral load; (B) baseline CD4 cell count

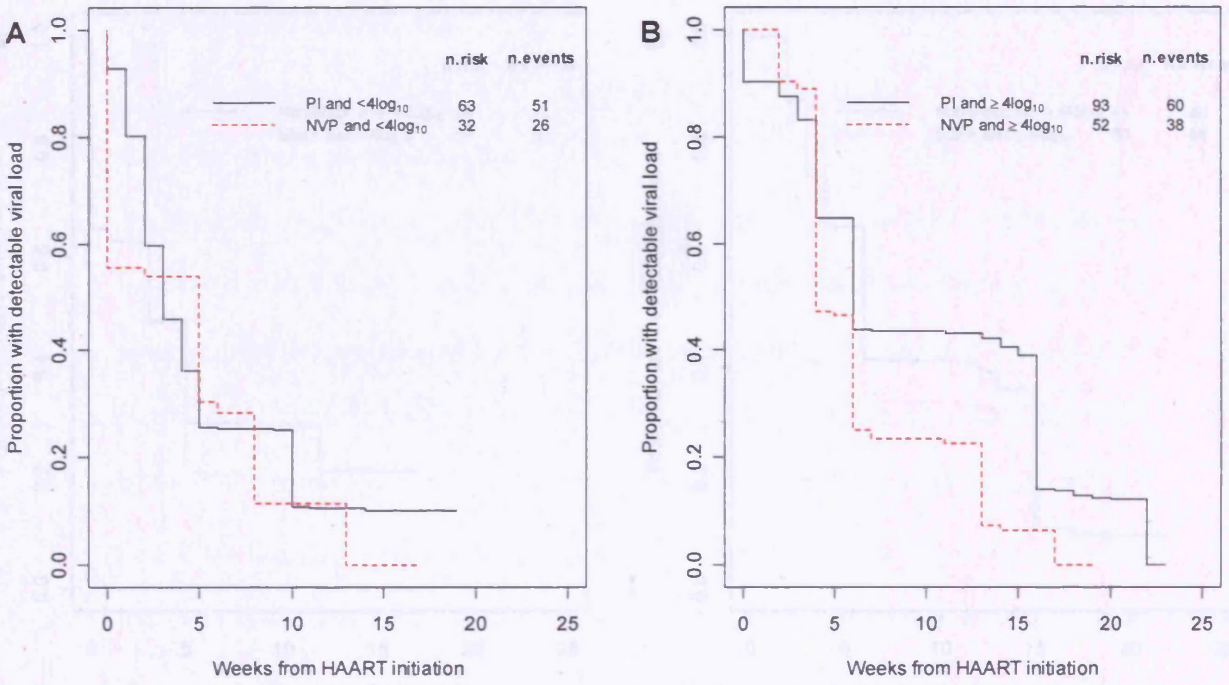


Weeks from initiation	*n* at risk (Fig 4.5A)			*n* at risk (Fig. 4.5B)		
	≥5 log_{10}	*4-4.99 log_{10}*	*<4 log_{10}*	*<200 cells/mm^3*	*200-499 cells/mm^3*	*≥500 cells/mm^3*
0	20	125	95	48	124	48
5	14	106	74	44	108	42
10	10	73	49	33	74	25
15	9	36	26	19	43	16
20	9	15	9	10	16	6
25	3	3	3	2	3	2

Figure 4.6A revealed overall similar proportions achieving undetectable HIV RNA levels after the time of initiation by HAART regimen for women with baseline HIV RNA <4 $\log_{10}$ copies/ml. Figure 4.6B revealed similar proportions by HAART regimen for women with baseline HIV RNA ≥4 $\log_{10}$ copies/ml in the first few weeks after initiation, after which point women receiving NVP-based HAART had a greater tendency to achieve undetectable levels. For women with baseline HIV RNA >4 $\log_{10}$ copies/ml, 35.2% (95%CI 20-51) and 53.0% (95%CI 20-94) of the PI and the NVP groups respectively reached undetectable levels by 5 weeks, 56.4% (95%CI 45-71) and 76.4% (95%CI 56-94) by 10 weeks and 59.4% (95%CI 43-75) and 93.4% (95%CI 66-99) by 15 weeks, highlighting a potential differential response by treatment category.

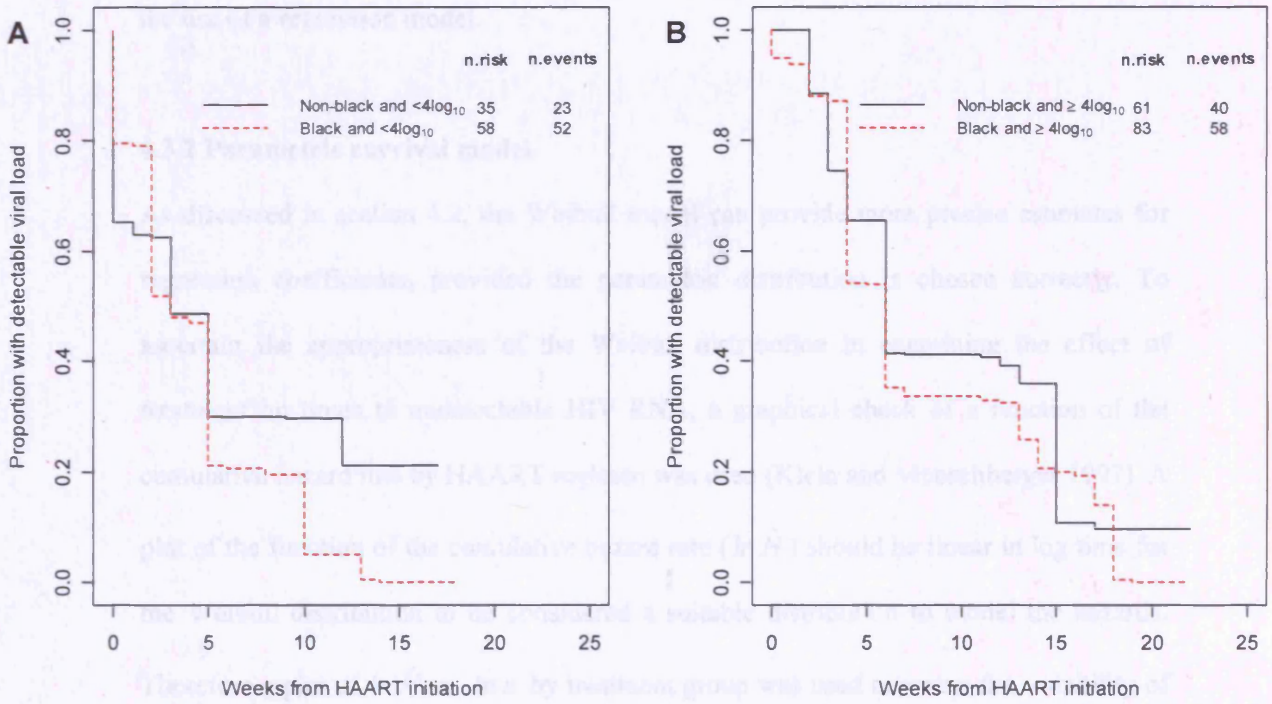
Figure 4.7 displays the proportions of women achieving undetectable viral load from time of initiation stratified by race and baseline viral load. The proportion reaching undetectable virus for non-black and black women with baseline HIV RNA ≥4 $\log_{10}$ copies/ml was 58.7% (95%CI 44-75) and 66.3% (95%CI 53-79) respectively at 10 weeks and 64.1% (95%CI 38-85) and 79.3% (95%CI 65-93) at 15 weeks, suggesting the possibility of race as a prognostic indicator (Figure 4.7B). Figure 4.7B also revealed a crossing of the two survival curves, supporting the use of a Weibull survival model, which does not require the assumption of proportional hazards to hold (Cox *et al.* 2007).

Figure 4.6 Survival curves for the time from initiation of HAART to achieving undetectable viral load by initial treatment category: women with (A) baseline viral load <4 log_{10} copies/ml; (B) baseline viral load ≥ log_{10} copies/ml



	n at risk (Fig. 4.6A)		*n* at risk (4.6B)	
Weeks from initiation	***PI***	***NVP***	***PI***	***NVP***
0	63	32	93	52
5	53	30	84	48
10	35	16	61	30
15	21	7	36	12
20	9	1	16	5
25	3	1	3	2

Figure 4.7 Survival curves for the time from initiation of HAART to achieving undetectable viral load by race category: women with (A) baseline viral load <4 log_{10} copies/ml; (B) baseline viral load ≥ log_{10} copies/ml



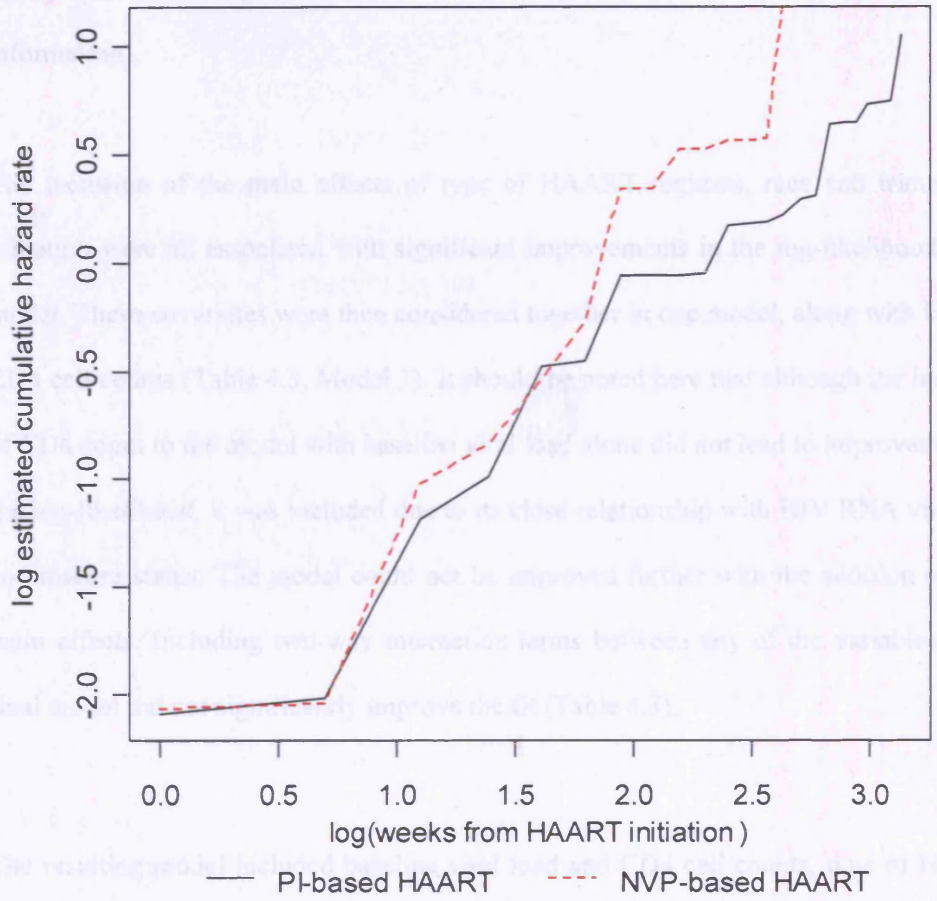
	n at risk (Fig. 4.7A)		*n* at risk (4.7B)	
Weeks from initiation	***Non-black***	***Black***	***Non-black***	***Black***
0	35	58	61	83
5	26	54	49	70
10	18	31	37	45
15	12	13	19	25
20	6	6	9	9
25	3	1	3	1

As noted in Section 4.2.2, the survival curves and estimates produced by the Turnbull method are purely for use as an explanatory tool, they do not take account of any other variables and therefore only very limited inferences can be made on these findings. The appropriate way to further assess any potential differences in survival curves is through the use of a regression model.

4.3.2 Parametric survival model

As discussed in section 4.2, the Weibull model can provide more precise estimates for regression coefficients, provided the parametric distribution is chosen correctly. To ascertain the appropriateness of the Weibull distribution in examining the effect of treatment on times to undetectable HIV RNA, a graphical check of a function of the cumulative hazard rate by HAART regimen was used (Klein and Moeschberger 1997). A plot of the function of the cumulative hazard rate ($\ln \hat{H}$) should be linear in log time for the Weibull distribution to be considered a suitable distribution to model the hazards. Therefore a plot of $\ln \hat{H}$ vs. $\ln x$ by treatment group was used to assess the suitability of the Weibull distribution here, with the log cumulative hazard rates estimated from the survival rates obtained from the Turnbull method. The resulting curves for the two treatment categories shown in Figure 4.8 are roughly linear for each group, except for some variation in the tail for the NVP-based group. This suggests that the Weibull parametric model provides an adequate fit for these groups.

Figure 4.8 Weibull hazard plot by HAART category.



Twenty-three women with information missing on race (3/23) or CD4 counts (20/23) were excluded; these women had similar characteristics with respect to baseline viral load (median 4.35 [IQR 3.88-4.77] $\log_{10}$ copies/ml for those with missing information vs. 4.14 [IQR 3.61-4.56] for those with complete information; Wilcoxon test p=0.19), baseline CD4 counts (median 246 [IQR 204-470] vs. 330 [IQR 210-480]; Wilcoxon test p=0.85). There were no significant differences between women with missing information and those with complete information with respect to the numbers receiving PI-based HAART (15/23: 65% vs. 160/217: 74%; χ^2=0.84, p=0.84), the numbers reaching undetectable viral load by delivery (15/23: 65% vs. 160/217: 74%; χ^2=0.39, p=0.53) and the numbers reaching undetectable levels by HAART regimen (60% [9/15] and 75%

[6/8]) by PI-based and NNRTI-based HAART respectively for women with missing information vs. 72% [102/141] and 76% [58/76], respectively for those with complete information).

The inclusion of the main effects of type of HAART regimen, race and trimester of initiation were all associated with significant improvements in the log-likelihood of the model. These covariates were then considered together in one model, along with baseline CD4 cell counts (Table 4.3, Model 1). It should be noted here that although the inclusion of CD4 count to the model with baseline viral load alone did not lead to improvements in the log-likelihood, it was included due to its close relationship with HIV RNA viral load and disease status. The model could not be improved further with the addition of other main effects. Including two-way interaction terms between any of the variables in the final model did not significantly improve the fit (Table 4.3).

The resulting model included baseline viral load and CD4 cell counts, type of HAART regimen, race and trimester of initiation. The RH for black race after adjusting for baseline viral load only (RH 1.47; 95%CI 1.04-2.07) and all other covariates in the model (RH 1.49; 95%CI 1.04-2.13) confirmed the differences seen in the survival curves in Figure 4.3B and 4.7. It should be noted here that the inclusion of IDU to the model neither improved the model fit in terms of the LRT, nor resulted in a significant change in the estimate of the black race estimate or its standard error and was therefore not included in the final model. As the majority of black women in this cohort were

Table 4.3 Likelihood ratio tests for model selection procedure

Model no.	Model	LRT	d.f.[a]	*p*
1	HAART regimen+ race+ VL+ CD4+ trimester of initiation	-	8	-
2	Model 1 + race*HAART	0.52	1	0.46
3	Model 1 + race*VL	4.27	2	0.12
4	Model 1 + race*CD4	5.49	2	0.06
5	Model 1 + race*trimester	0.92	2	0.63
6	Model 1 + VL*CD4	1.85	4	0.76
7	Model 1 + VL*HAART	0.19	2	0.91
8	Model 1 + CD4*HAART	2.13	2	0.34
9	Model 1 + VL*trimester	4.27	4	0.37
10	Model 1 + CD4*trimester	1.67	4	0.80

[a] - degrees of freedom

from different regions of Africa, these race-associated differences were explored in further models through the examination of region of birth, categorised as Non-African and Eastern, Central, Northern/Southern or Western Africa. Although constituting a small group of women, the Northern/Southern Africa group were included in a separate group as they could not feasibly be included in any of the other African regions and it was preferable to the alternative of removing them from the analysis completely; this however limits any inferences made on the resulting RH estimate. The presence of interaction terms in the model with region of birth instead of race was assessed; none were required. The univariable (adjusted for baseline viral load only) and multivariable RHs are shown in Table 4.4.

Table 4.4 Factors associated with time to achieving undetectable HIV RNA viral load after initiation of HAART in pregnancy - model without propensity score (n=217)

		Univariable analysis		Multivariable analysis	
	n	RH[a] (95% CI)	*p*	RH (95% CI)	*p*
Region of birth					
Non-African	96	1.00		1.00	
Eastern Africa	42	1.15 (0.73-1.83)	0.55	1.42 (0.86-2.32)	0.17
Central Africa	33	1.58 (0.96-2.60)	0.07	1.43 (0.84-2.42)	0.19
Northern/Southern Africa	7	0.82 (0.31-2.13)	0.68	1.00 (0.37-2.67)	0.99
Western Africa	39	1.58 (1.0-2.50)	0.05	1.81 (1.11-2.96)	0.02
HAART regimen					
PI-based	141	1.00		1.00	
NVP-based	76	1.62 (1.14-2.31)	<0.01	1.60 (1.10-2.33)	0.02
Time of initiation of HAART					
First trimester	12	1.00		1.00	
Second trimester	150	1.05 (0.55-2.04)	0.87	1.08 (0.54-2.17)	0.22
Third trimester	55	2.18 (1.0-4.74)	0.05	2.16 (0.95-4.88)	0.07
Baseline HIV RNA (copies/ml)[b]					
≥4.40 $\log_{10}$	72	1.00		1.00	
3.81-4.39 $\log_{10}$	73	1.75 (1.15-2.66)	<0.01	1.76 (1.13-2.75)	0.01
< 3.81 $\log_{10}$	72	2.54 (1.69-3.80)	<0.001	2.46 (1.57-3.83)	<0.001
Baseline CD4 count (cells/mm^3)					
<200	47	1.00		1.00	
200-499	123	1.39 (0.89-2.15)	0.14	1.41 (0.90-2.21)	0.14
≥500	47	1.42 (0.83-2.43)	0.20	1.52 (0.85-2.74)	0.16

[a] RH, Relative hazard, [b] Baseline HIV RNA categorised according to their tertiles

NOTE: Univariable and multivariable estimates adjusted for baseline viral load; multivariable estimates adjusted for all covariates listed in table

The probability (propensity score) of a woman being treated with NVP-based HAART was estimated using a logistic regression model with the following covariates; timing of first HIV diagnosis during pregnancy, region of birth, trimester of initiation, year of delivery and baseline viral load and CD4 cell counts. The unadjusted and adjusted odds ratios (OR and AOR, respectively) for the propensity score are given in Table 4.5. The inclusion of the timing of first HIV diagnosis was regarded a priori as an important consideration for clinicians in assigning a treatment to ARV-naïve pregnant women and was therefore included in the final model although it had a $p>0.5$. There was some evidence to suggest that women with baseline CD4 counts >200 cells/mm^3 had a higher odds of being treated with NVP-based HAART than women with counts <200 cells/mm^3, as did women delivering in the period 2003-2004, compared to women delivering in the period 1998-2000. Additionally, there was some evidence that women initiating HAART in the second trimester of pregnancy were less likely to receive NVP than women initiating earlier in pregnancy (Table 4.5). The propensity scores from the adjusted model were then estimated and a histogram of the scores is shown in Figure 4.9. The majority of the values were clustered between 25 and 50%; Therneau et al. describe only small alterations in the RH estimates when regression models are adjusted by scores clustered around this range (Therneau and Grambsch 2000).

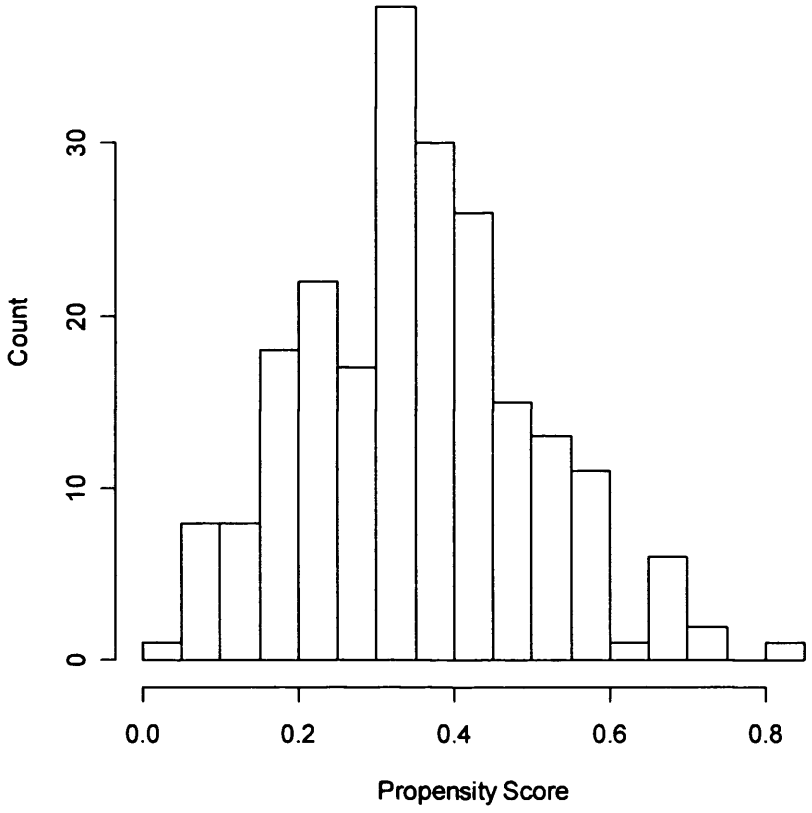
Table 4.5 Unadjusted and adjusted logistic regression models for the probability of being treated with NVP-based over PI-based HAART (*n*=217)

	Unadjusted		Adjusted	
	OR[a] (95% CI)	*p*	AOR[a, b] (95% CI)	*p*
Region of birth				
Non-African	1.00		1.00	
African	0.70 (0.40-1.22)	0.21	0.71 (0.39-1.32)	0.28
Timing of first HIV diagnosis				
Before pregnancy	1.00		1.00	
During pregnancy	1.30 (0.73-2.33)	0.37	1.09 (0.58-2.06)	0.79
Time of initiation of HAART				
First trimester	1.00		1.00	
Second trimester	0.31 (0.09-1.02)	0.05	0.26 (0.07-0.95)	0.04
Third trimester	0.55 (0.16-1.96)	0.36	0.46 (0.12-1.82)	0.27
Baseline HIV RNA (copies/ml)				
≥4.40 $\log_{10}$	1.00		1.00	
3.81-4.39 $\log_{10}$	1.44 (0.91-19.68)	0.07	4.04 (0.80-20.33)	0.09
< 3.81 $\log_{10}$	3.19 (0.68-15.11)	0.14	2.23 (0.42-11.86)	0.35
Baseline CD4 count (cells/mm^3)				
<200	1.00		1.00	
200-499	2.29 (1.04-5.03)	0.04	2.23 (1.01-5.36)	0.05
≥500	2.51 (1.01-6.23)	0.05	2.36 (0.85-6.57)	0.10
Time period				
1998-2000	1.00		1.00	
2001-2002	1.45 (0.68-3.09)	0.34	1.72 (0.76-3.90)	0.19
2003-2004	2.01 (0.91-4.44)	0.09	2.43 (1.03-5.76)	0.04

[a] – OR: Odds ratio; AOR: Adjusted Odds Ratio

[b] - Model adjusted for all covariates listed in table

Figure 4.9 Distribution of propensity score for treatment with NVP-based HAART



The estimated propensity scores were included in the model specified in Table 4.4 after stratification into quintiles. Adding the propensity score to the model had only a small effect on the resulting adjusted RH estimates and their standard errors. There was a marginal attenuation of the RH estimate for NVP-based HAART to 1.54 (95%CI 1.05-2.26), which could be a result of the reduction in bias through inclusion of the propensity score. Additionally, the estimated RH women for Western African origin was 1.90 (95%CI 1.16-3.12) compared to women of Non-African origin and baseline HIV RNA levels of 3.81-4.39 $\log_{10}$ and <3.81 $\log_{10}$ copies/ml had estimated RHs of 1.70 (95%CI 1.08-2.68) and 2.76 (95%CI 1.68-4.52), respectively. Although the resulting change in RH was only small, subsequent inferences were made and sensitivity analyses carried out on the adjusted survival model which included the propensity score

4.3.3 Model assessment

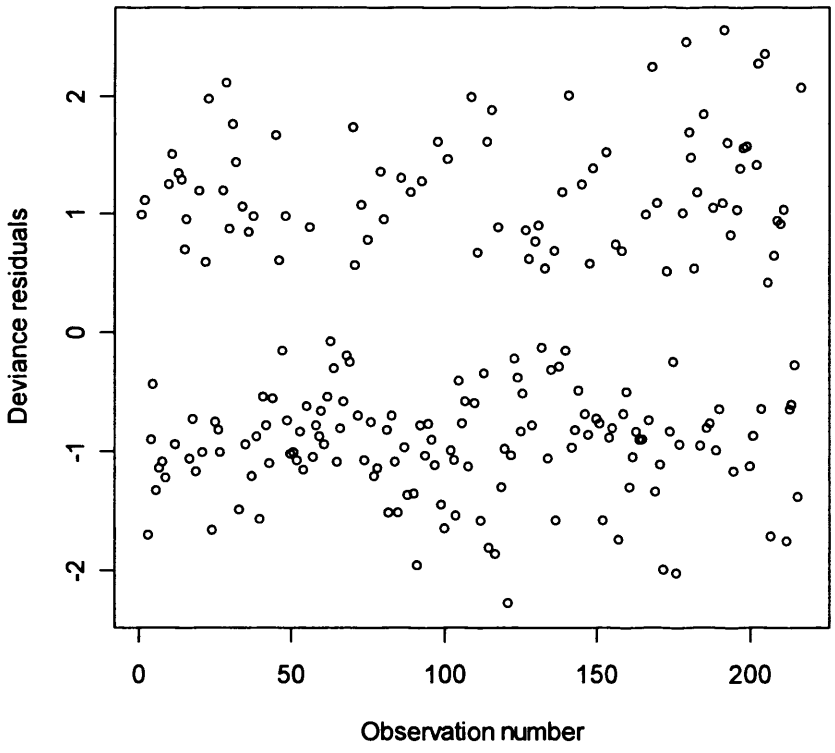
The scale parameter for the Weibull model, α, was estimated to be 1.43 (95%CI 1.10-1.63); the confidence interval for this estimate is bounded away from 1 and suggests that the exponential model would not be a simpler parametric alternative (Klein and Moeschberger 1997).

A plot of the deviance residuals against index number (i.e. against the index number for subjects in the model, 1,2,....,*n*, in no particular order) has been suggested as a guide to detect outliers from a Weibull model (Klein and Moeschberger 1997). A plot of the deviance residuals against observation number (i.e. 1, 2,...., 217) from the final survival model, which included the propensity scores is given in Figure 4.10 and revealed no indication of any major outliers.

4.3.4 Interpretation

The rate of women achieving undetectable viral load in the NVP group was estimated to be almost one and a half times of that in the PI group (ARH 1.54 [95%CI 1.05-2.26]). Women of Western African origin had an almost doubled rate of achieving undetectable viral load compared to women of Non-African origin (ARH 1.90 [95%CI 1.16-3.12]). Women with the lowest baseline viral loads (<3.81 $\log_{10}$ copies/ml) had the fastest rate of achieving undetectable viral load compared to women with baseline levels ≥4.40 $\log_{10}$ copies/ml (ARH 2.76 [95%CI 1.68-4.52]). There was no evidence of an association between baseline CD4 count and the rate of achieving undetectable vial load (RHs for baseline CD4 cell counts 200-499 and ≥500 cells/mm^3 were 1.25 [95%CI 0.69-2.24] and 1.40 [95%CI 0.71-2.76] compared to <200 cells/mm^3).

Figure 4.10 Deviance residuals from the adjusted Weibull model



Using the acceleration format parameterisation of the Weibull model, the median time to undetectable viral load for a woman receiving PI-containing HAART was estimated to be 1.38 (95% CI 1.04-1.83) times that of a woman receiving NVP-containing HAART. The acceleration factor for the model without the propensity score was 1.42 (95%CI 1.08-1.88), similar to that which included the propensity score. The predicted median time to undetectable virus for non-African women with baseline CD4 counts 200-499 cells/mm^3 and viral load 3.81-4.39 $\log_{10}$ copies/ml initiating treatment in the second trimester was 7.1 weeks (95% CI 3.60-10.53) for the NVP group and 9.8 weeks (95% CI 5.38-14.16) for the PI group; for Western African women with similar characteristics, these times were 4.4 weeks (95% CI 2.1-6.7) and 6 weeks (95% CI 3.2-8.9), respectively.

4.3.5 Sensitivity analysis

The majority of women receiving PI-based HAART received NFV, and therefore a sensitivity analysis to examine the time taken to achieve undetectable viral load between women receiving NFV-based HAART or NVP-based HAART was carried out. The resulting model revealed similar RHs for NVP-based HAART (RH 1.56, 95%CI 1.05-2.32). Similarly, a sensitivity analysis including only women initiating HAART in the first and second trimesters was carried out and resulted in an increase in the RH for NVP-based HAART (RH 1.82, 95%CI 1.18-2.84). The impact of including the baseline HIV RNA (log_{10} copies/ml) as a continuous covariate was also evaluated and did not lead to a substantial change in the relative hazard for NVP-based HAART (RH 1.50 95%CI 1.02-2.19).

Sub-analysis for women with CD4<250 cells/mm^3

The time to achieving undetectable virus was also examined for the 70 women who would be eligible for NVP-based HAART according to current prescribing advice (Public Health Service Taskforce 2006b) (i.e. with baseline CD4 counts <250 cells/mm^3). The median baseline viral load for this group of women was 4.35 log_{10} copies/ml (IQR 4.05-4.69) and the median time of initiation was 22 weeks gestation and 16 (24%) actually received NVP.

Using the Turnbull method, the estimated proportion reaching undetectable viral load at 5 weeks was 34.8% (95%CI 20.4-46.7) for those on PI and 52.9% (95%CI 22.1-71.6) for those receiving NVP, increasing to 50.4% (95%CI 34.0-62.8) and 82.4% (95%CI 50.7-93.7) at 8.5 weeks (due to small numbers of women, the Turnbull estimates and associated bootstrap confidence intervals were unstable after 8.5 weeks and were therefore, not reported here). In a Weibull regression model adjusting for race, timing of initiation in pregnancy (first/second trimester vs. third trimester), baseline viral load, and

CD4 cell count (<200 cells/mm^3 or between 200-250 cells/mm^3), the estimated RH for NVP-based vs. PI-based HAART was 1.89 (95%CI, 0.87-4.12). The RH was considerably larger than when women in all CD4 categories were included. However, although the power in this study may have been limited owing to the small sample size and this did not reach statistical significance, the set of values covered by the 95% CI were not clustered closely to 1 suggesting that a larger effect of NVP-based HAART among women with CD4 counts <250 cells/mm^3 should not be ruled out based on these data.

Women not achieving undetectable levels by the time of delivery

The 65 (27%) women delivering with detectable viral load were not significantly different to those achieving undetectable levels, with respect to race (31/64: 48% non-black vs. 110/173: 64% black; χ^2=3.84, p=0.05) and type of HAART (20/65: 31% PI-based vs. 64/175: 37% NVP-based HAART; χ^2 =0.47, p=0.49). However, more were severely immunosuppressed (CD4<200 cells/mm^3) (33% [19/58] vs. 18% [29/162], p=0.03) and more had baseline viral loads >5 $\log_{10}$ copies/ml (15% [10/65] vs. 6% [10/175], p=0.03). The median viral load at delivery for the 65 women not achieving undetectable viral load was 2.48 $\log_{10}$ copies/ml (IQR 2.26-3.11) and only 20 (31%) had viral load above 1000 copies/ml.

4.4 Key points

- Most (73%) ARV-naïve women initiating HAART antenatally delivered with undetectable virus and the remainder delivered with detectable, but generally low levels.

- Adjusting for baseline prognostic factors, the RH of achieving undetectable viral load was greater for women on NVP-containing HAART than for women on PI-based HAART, with women in the PI group taking on average 1.4 times longer to achieve undetectable levels.

- Women with lower baseline viral load and of Western African origin had a higher rate of achieving undetectable levels.

- There was no statistically significant differences in the time to attaining undetectable viral load by baseline CD4 status, although immunosuppressed women (CD4 count <200 cells/mm^3) had a non-significantly lower rate.

- A sensitivity analysis comparing NFV-based PI regimens only with NVP-based regimens confirmed a higher rate of achieving undetectable viral load with NVP.

- The use of an interval-censored survival model to take into account the non-uniform intervals of HIV RNA measurement in pregnancy should improve the accuracy of the treatment effect estimates.

- Adjustment for the propensity score for treatment resulted in only negligible changes in the treatment effect HR.

Chapter 5 Bivariate longitudinal model of HIV RNA and CD4 counts in pregnant women on HAART

5.1 Introduction

The probability of further live births in HIV-infected pregnant women has increased substantially in Europe since 1995 especially among black women, and the time interval between births has shortened (European Collaborative Study 2005a). An increasing proportion of HIV-infected pregnant women in Europe can therefore be expected to be on treatment at or before the time of conception.

Viral load in women receiving HAART at the time of conception would be expected to be low or at undetectable levels for those responding successfully to treatment. Women receiving HAART at conception with detectable viral load at their first visit in pregnancy are a very different group with a comparatively higher risk of transmitting their infection to their unborn child (European Collaborative Study 1999;Cooper *et al.* 2002) which may be an indication of a poorer response to treatment, non-compliance or resistance. Changes in viral load and CD4 count during pregnancy are of interest to evaluate the efficacy of treatment and the risk of MTCT for this group of women, and the factors which may affect these levels, including ethnicity and history of IDU and the type of HAART regimen being taken at conception (Chapters 3 and 4).

Pregnancy data on viral load and CD4 count dynamics in women becoming pregnant on HAART are lacking at present. It has been shown in non-pregnant adults that the joint modelling of viral load and CD4 count is important due to the intrinsic correlation

between them and results in improved parameter estimates for both markers (Thiébaut *et al.* 2002;Thiébaut *et al.* 2005). A bivariate random effects model is advantageous as it allows for direct interpretation of the relationship between the two markers without assuming the dependence of one marker in relation to the other (Thiébaut *et al.* 2002).

Using a bivariate linear mixed effects (BLME) model taking into account left-censoring due to the lower quantification limit of HIV RNA assays (Thiébaut *et al.* 2005;Thiébaut *et al.* 2003), levels of viral load and CD4 in HIV-infected HAART-treated women who become pregnant with detectable levels of viral load and the factors associated with these levels, were examined. Understanding how these levels change over pregnancy and identifying the factors which affect these will help to inform clinical decision-making in this group of women. The correlation between the markers estimated from the bivariate model will be assessed as well as the fixed and random effects estimates compared to those estimated from the univariate models for CD4 counts and viral load to ascertain whether this is an appropriate method for analysis of HIV markers in treated pregnant women. Univariate models for viral load and CD4 as a function of the other marker will also be assessed.

The joint modelling of HIV RNA and CD4 for untreated women or women starting HAART for the first time is of interest but was not possible for this study due to the few measurements available on both markers for untreated women and the different initiation times for women starting HAART for the first time in pregnancy (Chapter 4).

5.2 Methods specific to this analysis

5.2.1 Data

There were 597 HIV-infected pregnant women receiving HAART at conception. Two-hundred and fifty-five women with undetectable viral loads at their first visit in pregnancy were excluded as the use of a linear model is not appropriate to explain the complicated pattern of viral load markers in this group where rebounds may occur (Fitzgerald *et al.* 2002) and the inclusion of these women would not allow for a bivariate model to be fitted. An additional 180 women with neither viral load nor CD4 counts available were also excluded. The exclusion of these women is unavoidable due to missing data and may introduce potential biases to the analysis; thus the results need to be interpreted under this condition (see Chapter 7)

The analysis was restricted to the remaining 162 women receiving HAART at conception with a detectable viral load at their first measurement in pregnancy, and at least one CD4 count available in pregnancy. The selection criteria were chosen so that a bivariate model of CD4 and HIV RNA could be fitted, but limit any inferences which can be made to women with detectable viral loads at first visit in pregnancy only (see Chapter 7).

Variables considered in the model included gestational age (expressed as weeks to the time of delivery), a variable combining race and history of IDU (non-IDU white, non-IDU black and white IDU; there were no black IDUs here), HAART type (PI-containing vs. NNRTI-containing), time period (1998-1999, 2000-2001 and 2001-2006) and type of HIV RNA assay (Roche vs other). These variables were deemed important potential confounders as they are known to or have the potential to be related to levels of CD4

counts or viral load (European Collaborative Study 1999;Anastos *et al.* 2000;Saul *et al.* 2001;Touloumi *et al.* 2004) (Chapter 3 and 4). Due to the available number of observations, region of birth was not explored in this analysis as stratification by region of birth and IDU resulted in overly small groups of women, which did not facilitate model fitting and convergence for the left-censored HIV RNA model. To maximise the sample size of women in this analysis, timing of initiation was not considered in the initial models as only 109 women had information available on timing of treatment. A sensitivity analysis of women with information available on initiation of treatment (receipt of HAART for $\geq$ 6 months vs. < 6 months) was carried out on the final model to determine the impact of the inclusion of this information on the parameter estimates.

A $\log_{10}$ transformation of CD4 count was sufficient to satisfy the normality and homoscedasticity assumptions of linear mixed effect (LME) models (see Results section, Figure 5.9 and 5.13B). A $\log_{10}$ transformation was also preferred to the often used square root transformation as it allows for direct interpretation of the parameter estimates.

5.2.2 Statistical methods

Random effects models are an appropriate method to deal with repeated measures data (Laird and Ware 1982;Pinheiro and Bates 2000). However, there are additional methodological issues to take into account in the analysis of longitudinal data on HIV markers such as HIV RNA viral load and CD4 counts.

The inherent natures of HIV RNA viral load, which represent the number of actively replicating HIV, and CD4 counts which are the target of HIV, mean that the both are

intrinsically correlated. Univariate models to describe the levels of each marker, dependent on the other one is a possible approach but requires that measurements are available at the same time and with the same frequency. A bivariate random effects model allows the study of the relationship between these two HIV markers without this requirement (Goldstein 1989;Thiébaut *et al.* 2003). The bivariate model has also been shown result in a more appropriate estimation of the correlation between CD4 and viral load, providing a better fit to the data than two separate univariate models (Boscardin *et al.* 1998;Thiébaut *et al.* 2002;Thiébaut *et al.* 2005).

However, the study of levels of these two markers is further complicated by the sensitivity of assays used to detect HIV RNA viral load. Undetectable levels of HIV RNA viral load occur when the virus is below the lower quantification limit of these assays. Crude approaches to deal with these left-censored values such as imputation of half the value of the assay limit have been shown to lead to biased estimates of model parameters and their standard errors (Jacqmin-Gadda *et al.* 2000;Hughes 1999;Thiébaut and Jacqmin-Gadda 2004). An approach such as the conditional likelihood method proposed by Jacqmin-Gadda et al to estimate LME models with left-censored longitudinal data leads to comparatively smaller bias and mean squared errors of parameter estimates (Jacqmin-Gadda *et al.* 2000).

A similar approach based on conditional likelihoods and which can be implemented in SAS was used here (Thiébaut and Jacqmin-Gadda 2004). This approach has been shown to be comparable to the approach of Jaqmin-Gadda et al (Jacqmin-Gadda *et al.* 2000) and can be extended to the bivariate case.

Bivariate linear mixed effects model

The extension of the usual univariate LME model (Laird and Ware 1982;Pinheiro and Bates 2000) can be extended to the BLME model as follows (Thiébaut *et al.* 2003;Goldstein 1989). Let $Y_i = \begin{bmatrix} Y_i^1 \\ Y_i^2 \end{bmatrix}$ be the response vector for the subject i=1, ...,N, where Y_i^k is the n_k^i-vector of measurements for the HIV marker k (k=1,2). Similar to the LME model in Chapter 3, assuming that viral load measurements within each woman are representative of that woman's viral load trajectory in pregnancy, there is no requirement that the same measurement times are sampled or that the same number of measurements is made on each subject (Goldstein 1989). As longitudinal measurements in pregnancy are expected to be few in number, the accuracy of parameter estimates can be increased with the inclusion of subjects with only one CD4 measurement or one viral load measurement. However, a single CD4 or viral measurement will contribute to the estimate of the intercept but will add little to the estimation of the slope.

The BLME model can be written as follows:

$$Y_i = X_i\beta + Z_i\gamma_i + \varepsilon_i$$

with $\varepsilon_i \sim N(0, \Sigma_i)$ and $\gamma_i \sim N(0, G)$, and

$$\beta = \begin{bmatrix} \beta^1 \\ \beta^2 \end{bmatrix}, X_i = \begin{bmatrix} X_i^1 & 0 \\ 0 & X_i^2 \end{bmatrix}, Z_i = \begin{bmatrix} Z_i^1 & 0 \\ 0 & Z_i^2 \end{bmatrix} \text{ and } \varepsilon_i = \begin{bmatrix} \varepsilon_i^1 \\ \varepsilon_i^2 \end{bmatrix}.$$

For each marker k (k=1,2), X_i^k is a $n_i \times p^k$ design matrix of explanatory variables, β^k is a p^k-vector of fixed effects, Z_i^k is a $n_i \times q^k$ design matrix and γ_i^k is a q^k-vector of individual random effects with $q^k \leq p^k$. One of the assumptions made here is that the the residuals for each marker, ε_i^1 and ε_i^2, are not correlated. The covariance matrix of

measurement errors is therefore a diagonal matrix $\sum_i$ containing the elements $\sigma^2_{\varepsilon_1}$ and $\sigma^2_{\varepsilon_2}$, representing the measurement error for each marker.

The covariance matrix for the random effects can be represented as

$$G = \begin{bmatrix} G^1 & G^{12} \\ G^{21} & G^2 \end{bmatrix}.$$

This matrix can be divided into three parts: i) G^1, the matrix of variances and covariances of random effects for the first marker, ii) G^2, the matrix of variances and covariances of random effects for the second marker, and iii) G^{12} the matrix of covariances between the random effects of each marker. The correlation between the two markers is taken into account through this sub-matrix of covariances (Thiébaut *et al.* 2005).

LME model with left-censoring of HIV RNA viral load

When one or more viral load measurements have left-censored values, the likelihood for estimating the LME model needs to be modified. The method proposed by Thiebaut et al using the NLMIXED procedure in SAS, takes into account left-censored values when estimating model parameters by maximising a full likelihood, distinguishing the contribution of the observed measures and left-censored measures, and is given below (Thiébaut and Jacqmin-Gadda 2004).

For any of the Y_i's, assume that there exists an n_i^O-vector of observed responses Y_i^O and a n_i^C-vector of censored responses Y_i^C, where $n_i^O + n_i^C = n_i$. The likelihood function for the parameter vector θ is given by

$$L(\theta)=\prod_{i=1}^{N}\left[\int_{R_q}\left\{\prod_{j=1}^{n_{iO}} f_{Y_{ij}^{O}|\gamma_i}\left(Y_{ij}^{O}\mid\gamma_i=u\right)\right\}\times\left\{\prod_{j=n_{iO}+1}^{n_{iC}}\phi_{Y_{ij}^{C}|\gamma_i}\left(Y_{ij}^{C}\mid\gamma_i=u\right)\right\}\times f_{\gamma_i}(u)du_1du_2...du_q\right] \quad (5.1),$$

where $Y_{ij}$is the jth measure at the time t_{ij} (j= 1,..., n_i) for the subject i (i =1,..., N), $f_{Y_{ij}^{O}|\gamma_i}(.|.)$ is the univariate normal conditional density of the observed measure j in subject i given the random effects and $\phi_{Y_{ij}^{C}|\gamma_i}(.|.)$ is the univariate normal distribution function of the censored measure j in subject i given the random effects. The computation of this likelihood leads to the integration over the random effects, $u = u_1$, u_2, ..., u_q, i.e. it requires computation of a multiple integral of dimension q. In essence, rather than imputing midpoints for undetectable values of viral load, this method assumes instead that the left-censored values are part of a Gaussian distribution and estimates viral load values by completing this distribution using the likelihood method described above.

In practice, the calculation of a q-dimensional integral limits the number of random effects that can be feasibly included in the model; the method has been found to work well with up to four random effects (Thiébaut *et al.* 2003), but a larger number may lead to computational problems and needs to be considered here when fitting the bivariate LME model taking into account left-censoring.

Calculating standard errors for the correlation coefficients

The covariance parameters and asymptotic covariance matrix for the random effects terms were used to calculate the standard error for the correlation coefficients using the Delta method (Lynch M. and Walsh B. 1998).

$$\rho = \frac{\sigma_{ab}}{\sqrt{\sigma_a^2 \sigma_b^2}},$$

where ρ is the correlation coefficient, σ_a^2 and σ_b^2 are the variances of the two estimated random effects parameters a and b, respectively and σ_{ab} is the covariance between them. The delta method yields the following expression for the variance of ρ:

$$\text{var}(\rho) = \rho^2 \left[\frac{\text{var}(\sigma_a^2)}{(4\sigma_a^2)^2} + \frac{\text{var}(\sigma_b^2)}{(4\sigma_b^2)^2} + \frac{\text{var}(\sigma_{ab})}{\sigma_{ab}^2} + \frac{2\text{cov}(\sigma_a^2, \sigma_b^2)}{4\sigma_a^2 \sigma_b^2} - \frac{2\text{cov}(\sigma_a^2, \sigma_{ab})}{2\sigma_a^2 \sigma_{ab}} - \frac{2\text{cov}(\sigma_b^2, \sigma_{ab})}{2\sigma_b^2 \sigma_{ab}} \right],$$

and allows calculating 95% confidence intervals for ρ using a Normal approximation.

5.2.3 Model selection

Mean plots of $\log_{10}$ transformed viral load and natural scale CD4 count over pregnancy and supersmoothers were used to examine the changes in these markers over pregnancy.

The bivariate model takes into account the correlation between the two markers through the covariance of the random effects terms and there is no requirement for the univariate LME components for each marker to include the same covariates. The model selection process therefore began with univariate LME models for each marker. The modelling process is given in Figure 5.1, and shows the order in which the models were fitted and in which they will be presented in this chapter, the relationships between them and the names with which each model will be referred to throughout the chapter.

The first step was to determine the appropriate functional form of $\log_{10}$ HIV RNA and $\log_{10}$ CD4 changes over gestational age in separate univariate LME models with the use

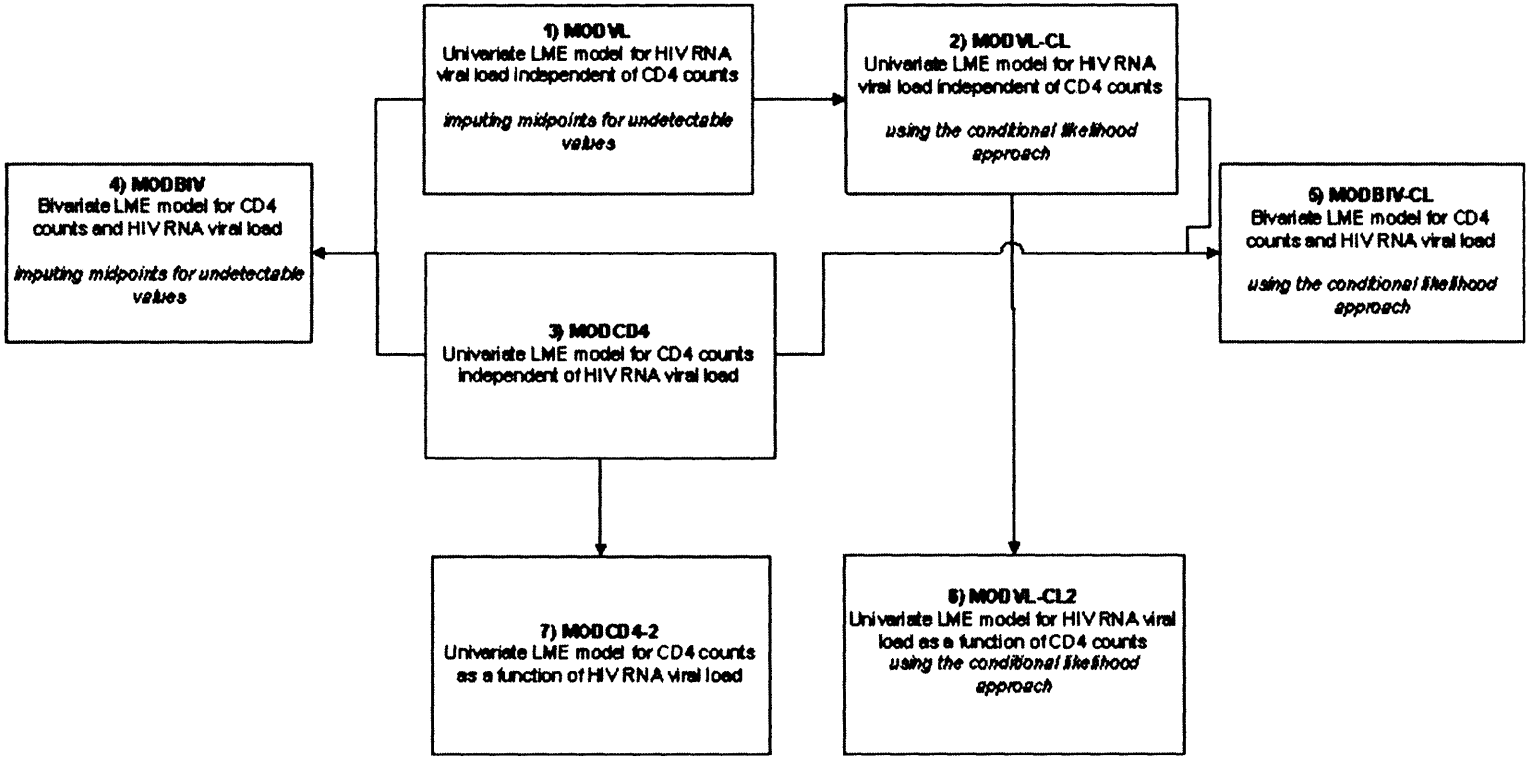
of Akaike's Information Criterion (AIC) and likelihood ratio tests (LRT), with midpoints imputed for undetectable viral load measurements. After determining the appropriate functional form between these markers and gestational age, the inclusion of additional covariates and interaction terms were tested using LRTs resulting in a univariate LME model for viral load, with imputed midpoints for undetectable values (MODVL) and a univariate LME model for CD4 counts (MODCD4).

Using the parameter estimates from MODVL as starting parameters for the iterative procedure, the model was refitted using the conditional likelihood approach assuming the same form for viral load changes over pregnancy and with the same covariates and random effects parameters (resulting in MODVL-CL).

The form and markers identified in univariate models (MODVL and MODCD4) were then included in a BLME model (MODBIV), dealing with the left-censoring using midpoints (Figure 5.1). The estimated correlation between the markers was assessed and the parameter estimates and their standard errors from this model were compared to those from the univariate LME models, MODVL and MODCD4.

Using the parameter estimates from MODBIV as starting parameters, the model was refitted using the conditional likelihood approach (MODBIV-CL). Finally, MODVL-CL was refitted with the inclusion of CD4 counts as a covariate (MODVL-CL2) and MODCD4 was refitted with the inclusion of HIV RNA viral load as a covariate (MODCD4-2).

Figure 5.1 Process used in modelling HIV RNA viral load and CD4 cell counts in pregnancy



An intention to treat approach was taken here; thus, interruptions and modifications in the treatment regimen were not taken into account in the analyses.

5.2.4 Model fitting and model selection

Model selection and estimation for MODVL, MODCD4 and MODCD4-2 were carried out using R version 2.4.1 (R Development Core Team 2006). The NLMIXED procedure in SAS was used to obtain MODVL-CL and MODVL-CL2. This procedure directly maximises the integrated likelihood in Equation 5.1 and the default method of the Quasi Newton-Raphson optimisation algorithm was used. Computation of the integral over the random effects was performed using an adaptive Gaussian quadrature method (SAS Institute Inc. 2004;Thiébaut and Jacqmin-Gadda 2004).

MODBIV was estimated with the SAS MIXED procedure and MODBIV-CL with the SAS NLMIXED procedure, with likelihood estimation and integral computations carried out using the above methods, unless otherwise stated (SAS Institute Inc. 2004).

An unstructured covariance matrix was assumed for random intercept and slope terms for all of the LME models estimated here. This allows estimating all the elements of this matrix without imposing any restrictions.

5.2.5 Model residuals

The residuals from MODCD4 and MODCD4-2 were examined using standard diagnostics for LME models (Pinheiro and Bates 2000). Assessing the residuals from MODVL-CL and MODVL-CL2 however, is complicated as the resulting residuals are right-censored. It has been suggested that a check of the model assumptions for such

residuals can be carried out by computing the cumulative frequency distribution of the residuals with a Kaplan-Meier estimator (Jacqmin-Gadda *et al.* 2000). The Kaplan-Meier estimate was therefore computed by replacing the survival time by the value of the residual, with the censoring indicator specifying whether the original value corresponding to the residual was left-censored or observed. This curve was then compared to the cumulative distribution of a normally distributed variable with variance σ_ε^2, i.e. the variance for the within-group residuals estimated from the viral load model.

One of the assumptions of the BLME model is that the (bivariate) residuals are not correlated, i.e. the within-group residuals for each response variable, ε^1 and ε^2 are independently and identically distributed (Gallant 1987;Goldstein 1989;Thiébaut *et al.* 2005). The Pearson correlation coefficient, and other scalar dependence measures, used to assess this relationship could be unreliable as the dependence will be averaged out over the range of the residuals. The local dependence map (LDM) (Jones and Koch 2003), an extension of the local dependence function (Holland and Wang 1987) was used to explore the dependence between bivariate residuals. The methodology for the estimation of the LDM is given in Appendix E.

5.3 Results

One hundred and sixty-two HIV-infected pregnant women were included in this analysis. The characteristics of these women are given in Table 5.1. Most women were white, although nearly a third were were black, 89% of whom were born in sub-Saharan Africa. Of the 100 women on a PI-based regimen at conception, 48 (48%) received nelfinavir, 17 (17%) indinavir, 14 (14%) lopinavir, ritonavir or saquinavir and the

Table 5.1 Characteristics of the 162 women receiving HAART at the time of initiation

Characteristic	Women	No. of measurements	
	n (%)	CD4	Viral load
		n=450	*n*=482
Type of HAART regimen			
PI-containing	100 (62)	268 (60)	280 (58)
NNRTI-containing	62 (38)	182 (40)	202 (42)
Race			
White	102 (67)	271 (63)	298 (64)
Black	44 (29)	136 (32)	140 (30)
Other	7 (4)	24 (5)	26 (6)
Unknown	9	19	18
Age at delivery (years)			
Median (IQR)	33 (30-36)		
18-25	18 (12)	42 (10)	48 (10)
26-34	79 (52)	231 (54)	256 (55)
≥35	55 (36)	154 (36)	160 (35)
Unknown	10	23	18
History of injecting drug use (IDU)			
Non-IDU	114 (72)	329 (74)	353 (74)
IDU	45 (28)	115 (26)	121 (26)
Unknown	3	6	8
Time period of delivery			
1998-1999	24 (15)	59 (13)	58 (12)
2000-2001	70 (43)	165 (37)	195 (40)
2001-2006	68 (42)	226 (50)	229 (48)
No. of previous live births			
None	59 (38)	154 (35)	164 (35)
1	59 (38)	179 (41)	195 (42)
2	27 (17)	73 (17)	76 (16)
≥3	11 (7)	30 (7)	34 (7)
Unknown	6	14	13
Timing of HAART (months before pregnancy)			
Median (IQR)	10.5 (4.8-21.8)	-	-
≥ 6 months	74 (68)	-	-
< 6 months	35 (32)	-	-
Unknown	54	-	-
HIV RNA assay			
Roche	-	-	375 (78)
Other	-	-	107 (22)

remaining 21 (21%) received boosted-PIs. Of the 62 women receiving NNRTI-based HAART, 53 (85%) received nevirapine and the remainder received efavirenz.

The median number of measurements per woman was 3 [IQR 2-3] for viral load and 3 [IQR 2-3] for CD4 counts; this limited any inferences on the shape of the patterns of HIV RNA and CD4 counts in pregnancy. Of the 482 viral load measures, 114 (24%) were undetectable (i.e. left-censored).

All women who had a history of IDU (28%) were white, of whom 6 (4%) actively used injecting drugs during pregnancy.

Table 5.2 shows the median CD4 counts and viral loads (with midpoints imputed for undetectable values) according to type of HAART regimen and by ethnicity and history of IDU, for the 143 women with complete information on ethnicity and IDU. This table includes all the CD4 and HIV RNA measurements available at these four time periods, ignoring the correlation between measurements on an individual and is used here as a crude exploratory tool for examining changes in the markers over pregnancy by the main covariates of interest.

There were 101 measurements (49 CD4 and 52 viral load measurements) on 19 women with missing information on ethnicity. The distribution of viral load for these 19 women was not significantly different to the 143 women with complete information (median 3.02 $\log_{10}$ copies/ml [IQR 2-3.71] vs. median 2.69 [IQR 2-3.48] $\log_{10}$ copies/ml; Wilcoxon test p=0.86). There were also no differences in the distribution of CD4 counts (median 408 cells/mm^3 [IQR 313-609] for women with incomplete information, vs.

Table 5.2 Median CD4 count and log_{10} viral load over pregnancy stratified according to type of HAART regimen, ethnic group and history of IDU (n=143)

	1st trimester	2nd trimester	3rd trimester	Delivery
ALL				
Number of women	74	114	99	61
CD4 count	381 (257-490)	376 (262-513)	447 (310-604)	407 (299-572)
Viral load	3.52 (2.39-4.43)	2.88 (2.30-3.55)	2.13 (1.40-2.93)	2.30 (1.60-2.64)
PI-based HAART				
Number of women	44	69	55	39
CD4 count	396 (280-502)	390 (282-522)	451 (329-601)	407 (297-514)
Viral load	3.60 (2.72-4.47)	2.86 (2.46-3.50)	2.30 (1.40-2.89)	2.40 (2.00-2.75)
NNRTI-based HAART				
Number of women	30	45	44	22
CD4 count	359 (252-460)	338 (252-491)	392 (301-619)	448 (309-669)
Viral load	3.30 (2.59-4.27)	2.89 (2.00-3.63)	2.00 (1.40-3.02)	2.00 (1.40-2.30)
Non-IDU White				
Number of women	35	42	44	22
CD4 count	424 (323-615)	365 (287-539)	469 (360-658)	451 (255-441)
Viral load	3.69 (2.66-4.60)	2.90 (2.30-3.64)	2.31 (1.40-3.18)	2.51 (2.00-2.87)
Non-IDU Black				
Number of women	22	39	27	25
CD4 count	313 (229-385)	290 (209-410)	309 (265-432)	290 (255-441)
Viral load	3.92 (2.79-4.43)	2.90 (2.33-3.55)	2.00 (1.40-2.64)	1.99 (1.40-2.40)
IDU White				
Number of women	17	33	28	14
CD4 count	400 (285-473)	476 (390-610)	493 (420-632)	552 (403-870)
Viral load	3.11 (2.61-3.33)	2.79 (2.08-3.43)	2.00 (1.60-3.18)	2.13 (2.00-2.46)

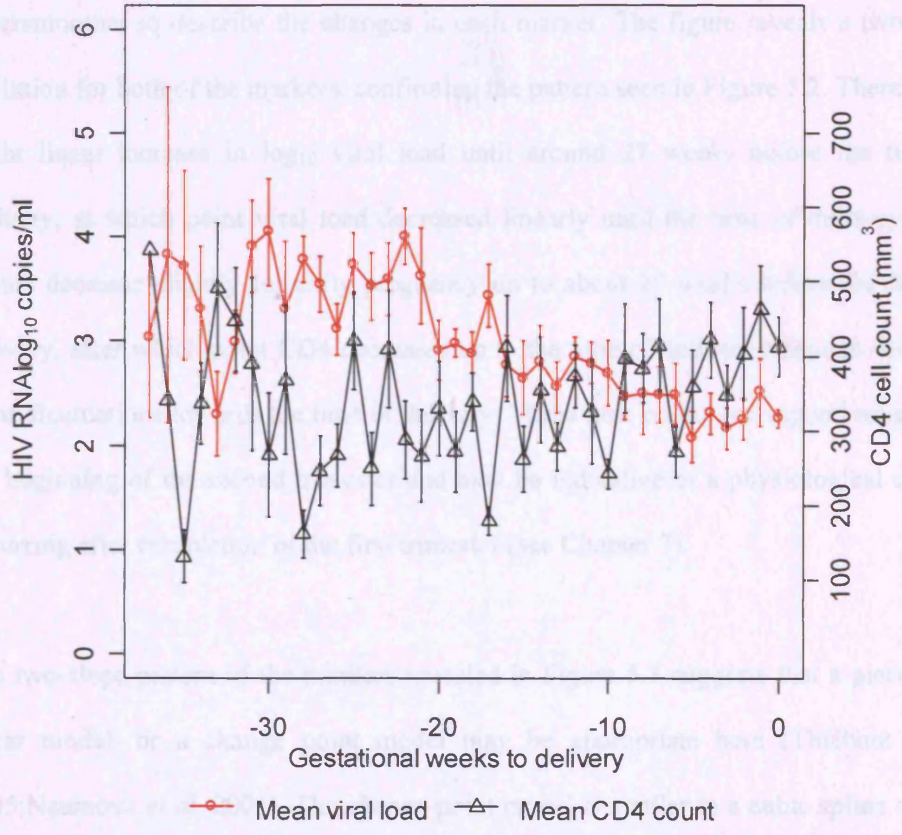
median 399 [IQR 281-545]; Wilcoxon test p=0.40). The principal aim of this analysis was to assess factors affecting the levels of CD4 and viral load over pregnancy, and these 19 women were excluded from subsequent analyses.

The median CD4 cell counts overall were similar in the first and second trimester but were comparatively higher in the third trimester and at the time of delivery (Table 5.2). Although there were no significant differences in the distribution of CD4 count between treatment group at the four time points shown (Wilcoxon test *p*-value 0.39, 0.32, 0.63 and 0.42 respectively; Table 5.2), counts were lower in the NNRTI group at all time periods except delivery. CD4 count values were lower in black women than in the other two groups and higher in white women with a history of IDU compared to other two groups for all four time points shown (Kruskal Wallis test for differences between the distributions for the three groups; *p*-value 0.04 for the first trimester and <0.01 for all other time points).

Median viral load decreased over the four time points and was lowest in the third trimester and at delivery. The number of undetectable values increased over pregnancy; 0% (0/81) in the first trimester, 18% (29/163) in the second trimester, 38% (47/125) in the third trimester and 46% (28/61) at the time of delivery. This is partly an artefact of the inclusion criteria of a detectable viral load at first visit in pregnancy. There were no differences in the distribution of viral load by type of HAART regimen in the first, second or third trimesters of pregnancy (Wilcoxon test *p*-values 0.48, 0.55 and 0.55, respectively); the test could not be carried out at delivery due to the large number of ties in the data. Viral load in the first and second trimester tended to be lower in white IDUs compared to white non-IDUs and black women, but overall the groups were similar with respect to their distributions of viral load at the four time points (Kruskal-Wallis test *p*-values 0.17, 0.68, 0.39 and 0.08 respectively; Table 5.2).

Figure 5.2 shows the weekly mean observed $\log_{10}$ viral load and natural scale CD4 counts over pregnancy with associated 95% confidence intervals for the 143 women.

Figure 5.2 Mean observed viral load and CD4 count during pregnancy for women becoming pregnant while receiving HAART (*N*=143). Bars show 95% confidence intervals



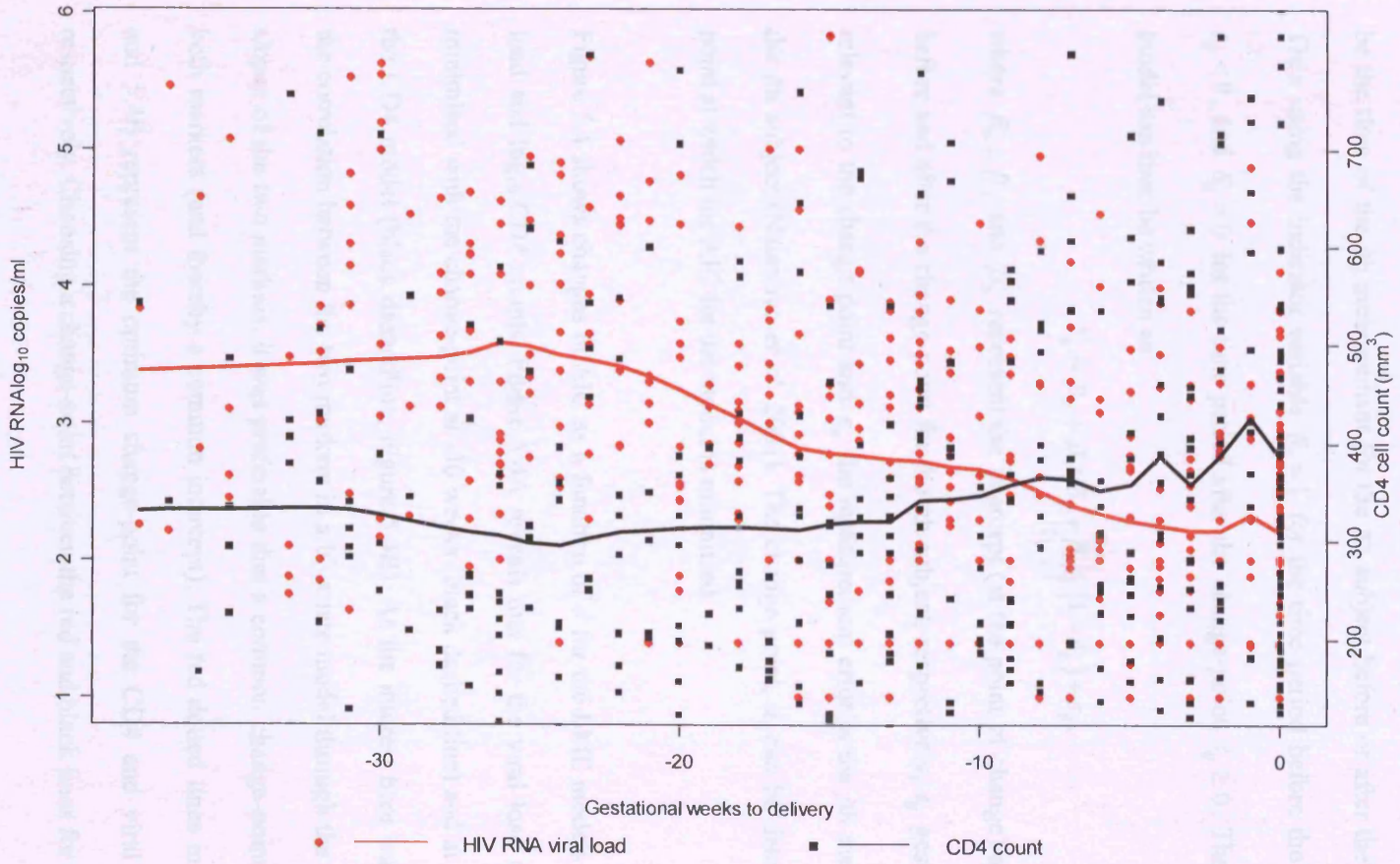
The narrowing of the confidence intervals over pregnancy reveals the larger number of measurements available in later pregnancy. This figure reveals some increases in viral load until around 30 weeks before delivery, at which point the mean viral load decreases up to the time of delivery. Mean CD4 counts can be seen to decrease up to around 28 weeks before the time of delivery, at which point mean CD4 is seen to decrease until the

time of delivery. This crude graphical approach describes the rough trend in the markers, and suggests that a change-point model may be suitable for both of them.

In order to smooth these changes over pregnancy to reveal the underlying structure of the $\log_{10}$ transformed HIV RNA viral load and CD4 count, supersmoothers were used. Figure 5.3 shows a scatterplot of viral load and CD4 count over pregnancy, with a supersmoother to describe the changes in each marker. The figure reveals a two-slope evolution for both of the markers, confirming the pattern seen in Figure 5.2. There was a slight linear increase in $\log_{10}$ viral load until around 27 weeks before the time of delivery, at which point viral load decreased linearly until the time of delivery. CD4 counts decrease slightly for early pregnancy up to about 25 weeks before the time of delivery, after which point CD4 decreases up to the time of delivery although there are some fluctuations towards the time of delivery. These time points correspond roughly to the beginning of the second trimester and may be indicative of a physiological change occurring after completion of the first trimester (see Chapter 7).

The two-slope pattern of the markers revealed in Figure 5.3 suggests that a piece-wise linear model, or a change point model may be appropriate here (Thiébaut *et al.* 2005;Naumova *et al.* 2001). The change-point model is similar to a cubic spline model, but the form of the model is constrained to be linear on either side of the point of change, or knot and is fitted as follows.

Figure 5.3 Supersmoothers for log10 viral load and CD4 count over pregnancy



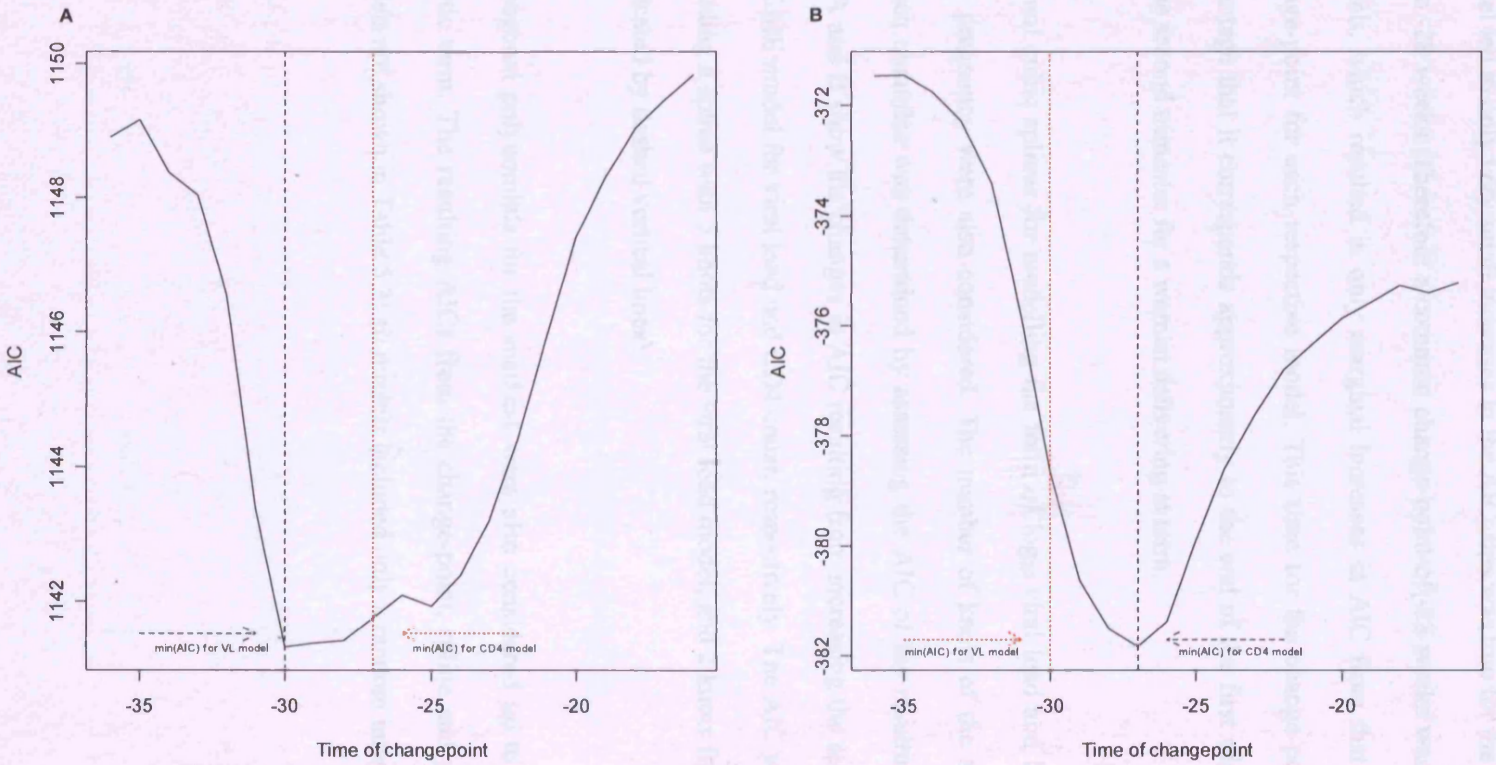
Let d represent the time of change-point for gestational weeks to delivery and g_{ij} gestational weeks to delivery at the jth measurement for the ith subject. Let $t_{ij} = g_{ij} - d$ be the time of the jth measurement for the ith subject before or after the change-point. Then using the indicator variable $\delta_{ij} = 1$ for the time period before the change-point, $t_{ij} < 0$, and $\delta_{ij} = 0$ for the time period after the change-point, $t_{ij} \geq 0$. The change-point model can then be written as

$$Y_{ij} = \beta_{0i} + \beta_{1i} t_{ij} \delta_{ij} + \beta_{2i} t_{ij} \left(1 - \delta_{ij}\right) + \varepsilon_{ij},$$

where β_{0i}, β_{1i} and β_{2i} represent the intercept (at the point of change) and two slopes before and after the change-point for the ith subject, respectively, t_{ij} gestational weeks relevant to the change point and ε_{ij} the measurement error at the jth measurement for the ith subject (Naumova *et al.* 2001). The change-point, d, can be determined as the point at which the AIC for the model is minimised.

Figure 5.4 shows changes in AIC as a function of d for the LME models for $\log_{10}$ viral load and $\log_{10}$ CD4 counts. Figure 5.4A reveals that for the viral load model, AIC is minimised with the change-point at -30 weeks (black dashed line) and at -27 weeks for the CD4 model (black dashedline, Figure 5.4B). As the interest here was in assessing the correlation between the two markers in a bivariate model through the intercepts and slopes of the two markers, it was preferable that a common change-point is chosen for both markers (and thereby a common intercept). The red dotted lines in Figures 5.4A and 5.4B represent the optimum change-point for the CD4 and viral load models, respectively. Choosing a change-point between the red and black lines for the viral load

Figure 5.4 Change in AIC from varying time of change-point for LME model for: A) viral load , B) CD4 count.



model led to only very small increases in the AIC; this was true for the CD4 model only up to -28 weeks. Therefore a common change-point of -28 weeks was chosen for both models, which resulted in only marginal increases in AIC from that of the optimum change-point for each respective model. This time for the change-point also had the advantage that it corresponds approximately to the end of the first trimester/beginning of the second trimester for a woman delivering at term.

Natural cubic splines for modelling the form of $\log_{10}$ viral load and $\log_{10}$ CD4 counts over pregnancy were also considered. The number of knots of the spline at suitably chosen quantiles was determined by assessing the AIC of the resulting model. Figure 5.5 A and B show the changes in AIC resulting from increasing the number of knots in the LME model for viral load and CD4 count, respectively. The AIC was minimised by including a spline with 5 knots for the viral load model, and 2 knots for the CD4 model (indicated by dashed vertical lines).

Orthogonal polynomials for the markers were also considered up to and including a quartic term. The resulting AICs from the change-point, spline and polynomial LME models are shown in Table 5.3; all models included only a random intercept term.

Figure 5.5 Changes in AIC from additional knots of the spline for: A) viral load, B) CD4 count model.

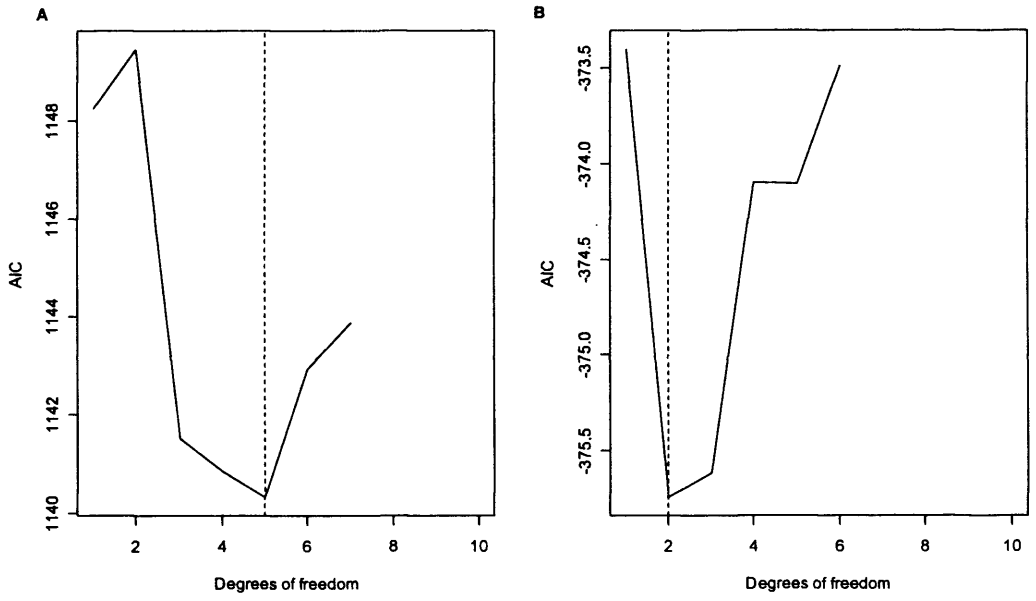


Table 5.3 AIC from models comparing functional forms for $\log_{10}$ viral load and $\log_{10}$ CD4 count

	AIC from LME model	
Model	**$\log_{10}$ Viral load**	**$\log_{10}$ CD4 count**
Change-point	1141.40	-381.52
Linear	1148.24	-373.40
Quadratic	1149.06	-376.18
Cubic	1139.46	-376.60
Quartic	1141.35	-374.62
Cubic spline	1140.32	-375.74

The change-point model resulted in the smallest AIC for modelling $\log_{10}$ CD4 counts over pregnancy, while the cubic model resulted in the smallest AIC for $\log_{10}$ viral load

(Table 5.3). However, little was gained in terms of the reduction in AIC (and therefore model goodness-of-fit) with the cubic over the change-point model for viral load. Additionally, the parameter estimates from an orthogonal polynomial model are difficult to interpret directly and therefore the change-point model was used for both markers, with a common change-point fixed at -28 weeks.

The inclusion of an additional random effect term for the first slope did not result in an improved log-likelihood for either the viral load model ($\chi^2_{df=2} = 3.12$; p=0.21) or the CD4 model ($\chi^2_{df=2} = 2.66$; p=0.26). However, inclusion of a random effect for the second slope resulted in a considerable improvement in the model for CD4 count ($\chi^2_{df=2} = 7.09$; p=0.03), and a relatively smaller improvement for the viral load model ($\chi^2_{df=2} = 4.51$; p=0.10). The random effect for the second slope was however retained for both models, so that the correlation between the two slopes (through their estimated covariance) could be investigated in the bivariate model, in addition to the correlation between the intercepts.

5.3.1 Model selection for univariate LME models (MODVL and MODCD4)

The main covariates of interest were considered in the change-point model for both markers separately and the resulting changes in log-likelihood and associated p-value from the LRTs are given in Table 5.4.

Table 5.4 Likelihood ratio tests for including main covariates

Covariates	Viral load model		CD4 count model	
	LRT $\chi^2_{(df)}$	p	LRT $\chi^2_{(df)}$	p
Race by IDU	5.06 $_{(2)}$	0.078	15.70 $_{(2)}$	<0.001
Type of HAART regimen	1.93 $_{(1)}$	0.16	0.075 $_{(1)}$	0.78
Time period	6.02 $_{(2)}$	0.049	2.100 $_{(2)}$	0.35
Type of assay	3.14 $_{(1)}$	0.076	-	-
Model 1 for viral load: Race/IDU, year and assay			-	-
Model 1- race/IDU	5.30 $_{(2)}$	0.06	-	-
Model 1 – type of assay	2.51 $_{(1)}$	0.11	-	-
Model 1 – time period	5.70 $_{(2)}$	0.06	-	-
Model 2 for CD4 count: Race by IDU only	-	-		
Model 2 + time period	-	-	0.260 $_{(2)}$	0.87
Model 2 + type HAART regimen	-	-	0.003 $_{(1)}$	0.95

CD4 model

Only the inclusion of the race by IDU history covariate resulted in an improvement in the CD4 change-point model (Table 5.4). The inclusion of an interaction term between the race/IDU covariate and either of the slopes did not result in an improvement in the fit of the model.

Viral load model

The inclusion of the main effects of race/IDU, time period or type of assay to the change-point model resulted in an improvement in the log-likelihood, and these were considered together in one model (Model 1, Table 5.4). Removal of the covariate for

Table 5.4 Likelihood ratio tests for including main covariates

Covariates	Viral load model		CD4 count model	
	LRT $\chi^2_{(df)}$	*p*	LRT $\chi^2_{(df)}$	*p*
Race by IDU	5.06 $_{(2)}$	0.078	15.70 $_{(2)}$	<0.001
Type of HAART regimen	1.93 $_{(1)}$	0.16	0.075 $_{(1)}$	0.78
Time period	6.02 $_{(2)}$	0.049	2.100 $_{(2)}$	0.35
Type of assay	3.14 $_{(1)}$	0.076	-	-
Model 1 for viral load: Race/IDU, year and assay			-	-
Model 1- race/IDU	5.30 $_{(2)}$	0.06	-	-
Model 1 – type of assay	2.51 $_{(1)}$	0.11	-	-
Model 1 – time period	5.70 $_{(2)}$	0.06	-	-
Model 2 for CD4 count: Race by IDU only	-	-		
Model 2 + time period	-	-	0.260 $_{(2)}$	0.87
Model 2 + type HAART regimen	-	-	0.003 $_{(1)}$	0.95

CD4 model

Only the inclusion of the race by IDU history covariate resulted in an improvement in the CD4 change-point model (Table 5.4). The inclusion of an interaction term between the race/IDU covariate and either of the slopes did not result in an improvement in the fit of the model.

Viral load model

The inclusion of the main effects of race/IDU, time period or type of assay to the change-point model resulted in an improvement in the log-likelihood, and these were considered together in one model (Model 1, Table 5.4). Removal of the covariate for

type of assay did not result in a significant worsening of the model, whereas the removal of either the covariates for race/IDU, or time period did. Therefore, the final model for viral load included the main effects of race/IDU and time period (MODVL). The inclusion of any interaction terms between the main effects of race/IDU and time period or between these covariates and either of the two slopes did not result in an improvement in the model.

The viral load model was fitted again with the same covariates and random effect terms, but using the conditional likelihood method which allows for left-censoring. The estimates from MODVL were used as starting parameters for the iterative procedure. However, the adaptive Gaussian quadrature method failed to compute the integral over the random effects. Changing the number of quadrature points manually and changing the integral method did not result in successful computation of the integral. Parameterisation of gestational weeks to the time of delivery as years instead of weeks was considered instead; this parameterisation increased the magnitude of the estimated slope and the magnitude of the associated random effects term for the slope. Fitting the model again using the adaptive Gaussian quadrature method with this parameterisation resulted in the successful computation of the integral and the model converged. This highlights the complex nature of this model and the importance of choosing both appropriate time-scales and starting parameters for successful estimation. The adjusted coefficients estimated from the two methods are given in Table 5.5, with the time of measurement reparameterised as weeks from delivery.

Table 5.5 Adjusted coefficients from MODVL and MODVL-CL (n=143)

	MODVL		*MODVL-CL*	
	log_{10} viral load		**log_{10} viral load**	
	Coefficient (95%CI)	***p***	**Coefficient (95%CI)**	***p***
Mean at 28 weeks before delivery	4.12 (3.70, 4.53)	<0.001	4.12 (3.63-4.61)	<0.001
Initial slope (<-28 weeks)	0.039 (-0.03, 0.10)	0.240	0.060 (-0.012, 0.13)	0.100
Slope to delivery (>-28 weeks)	-0.055 (-0.06, -0.045)	<0.001	-0.069 (-0.082, -0.056)	<0.001
Race by IDU				
White non-IDU	0.00		0.00	
Black non-IDU	-0.21 (-0.49, 0.063)	0.240	-0.15 (-0.49, 0.19)	0.370
White IDU	-0.31 (-0.59, -0.032)	0.030	-0.37 (-0.70, -0.024)	0.036
Time period				
1998-1999	0.00		0.00	
2000-2001	-0.24 (-0.64, 0.16)	0.250	-0.10 (-0.59, 0.39)	0.690
2002-2006	-0.44 (-0.82, -0.059)	0.025	-0.37 (-0.84, 0.092)	0.120
Random effects parameters				
Intercept	0.69 (0.50, 0.93)		0.62 (0.27, 0.84)	
Slope to delivery	0.02 (0.013, 0.042)		0.036 (0.013, 0.049)	
Correlation between intercept and slope	-0.70 (-0.88, -0.36)		-0.39 (-0.82, 0.05)	
Within-group residual deviation	0.74 (0.67, 0.82)		0.81 (0.71, 0.90)	

There was little change in the estimates of the mean viral load at the time of the change-point and their standard errors between the two models (as evidenced by the similar 95% CIs). There was some change in the estimated coefficients for the race covariates, with white women with a history of IDU estimated to have a slightly lower viral load when left-censoring was taken into account. Additionally, the estimated viral load for women delivering after 2001 was no longer significantly different from that of women delivering between 1998 and 1999, partly due to the larger estimated standard error from the conditional likelihood method. However, the parameter estimates suggested lower HIV RNA levels in later time periods, which could be representative of a general improvement in adherence and health of pregnant women over the years (the number of undetectable measurements increased over the three time periods, from 14% [5/36} in 1998-1999, to 17% [25/145] in 2000-2001 to 27% [55/206] in 2002-2006; $\chi^2_{trend} = 5.59$; p=0.018).

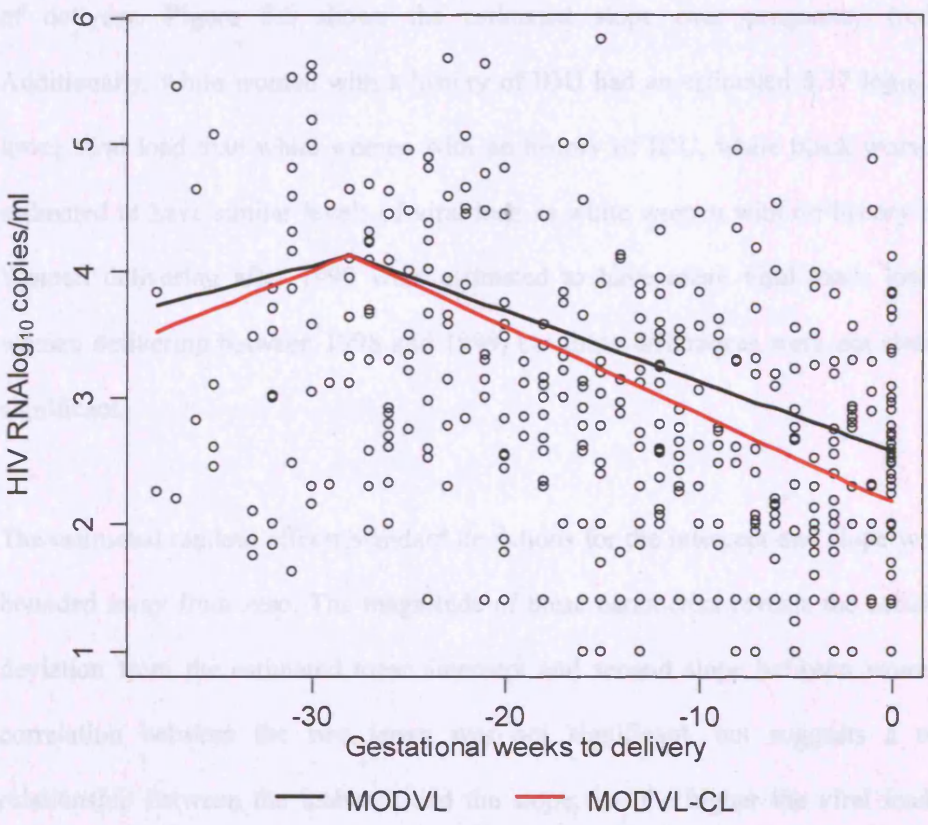
The main difference between the models however was in the estimation of the slopes. The estimate for the initial slope from MODVL-CL was greater than that from MODVL, although the 95% CI remained bounded by zero. The estimate of the second slope to delivery was also greater in MODVL-CL and was estimated to be almost 25% lower than that estimated with the simpler approach. Figure 5.6 shows the estimated slopes, highlighting the differences in the estimated slopes resulting from the two methods.

The estimated random intercept terms were similar between the two models; however those for the slopes were not. MODVL-CL estimated the random deviation around the slopes to be almost twice as large as that from the simpler model. Subsequently, the

correlation between the two intercepts was estimated to be almost half as small in MODVL-CL as in MODVL, and was not significantly different from zero.

MODVL-CL which accounts for the left-censored viral load using the more complex method is the preferred model, especially where interest lies in estimating the intensity of the slope of viral load.

Figure 5.6 Estimated mean change in log_{10} viral load accounting for undetectable values using the two approaches



5.3.2 Interpretation of univariate viral load with conditional likelihood approach (MODVL-CL)

Allowing for a change in slope of viral load at 28 weeks before delivery (roughly corresponding to the start of the second trimester), viral load was estimated to increase by 0.060 $\log_{10}$ copies/ml per week (95%CI -0.012, 0.13) in the period up to 28 weeks from delivery; this increase was not however significantly different from zero, i.e. the slope was stable (Table 5.5). Viral load thereafter was estimated to decrease by 0.069 $\log_{10}$ copies/ml per week (95%CI -0.082, -0.056) until the time of delivery; this corresponds to a 15% weekly decrease in copies/ml (95%CI -12%, -17%) until the time of delivery. Figure 5.6 shows the estimated slope over pregnancy (red line). Additionally, white women with a history of IDU had an estimated 0.37 $\log_{10}$ or 57% lower viral load than white women with no history of IDU, while black women were estimated to have similar levels of viral load as white women with no history of IDU. Women delivering after 1999 were estimated to have mean viral loads lower than women delivering between 1998 and 1999, but these differences were not statistically significant.

The estimated random effects standard deviations for the intercept and slope were both bounded away from zero. The magnitude of these parameters reveals the considerable deviation from the estimated mean intercept and second slope between women. The correlation between the two terms was not significant, but suggests a negative relationship between the intercept and the slope; i.e. the higher the viral load at the change point (28 weeks before delivery), the greater the decrease in viral load up to the time of delivery ($\hat{\rho} = -0.39$).

5.3.3 Univariate CD4 model and interpretation (MODCD4)

The table below shows the estimated coefficients from the adjusted univariate CD4 model.

Table 5.6 Adjusted coefficients from MODCD4 (n=143)

	$\log_{10}$ CD4 count	
	Coefficient (95%CI)	***p***
Mean at -28 weeks	2.56 (2.51, 2.61)	<0.001
Initial slope (<-28 weeks)	-0.013 (-0.02, -0.003)	0.008
Slope to delivery (>-28 weeks)	0.0037 (0.002, 0.005)	<0.001
Race by IDU		
White non-IDU	0.00	
Black non-IDU	-0.12 (-0.19, -0.05)	0.001
White IDU	0.028 (-0.05, 0.10)	0.46
Random effects parameters		
Intercept	0.18 (0.15, 0.22)	
Slope to delivery	0.004 (0.0025, 0.0062)	
Correlation between intercept and slope	-0.37 (-0.62, -0.051)	
Within-group residual deviation	0.09 (0.081, 0.10)	

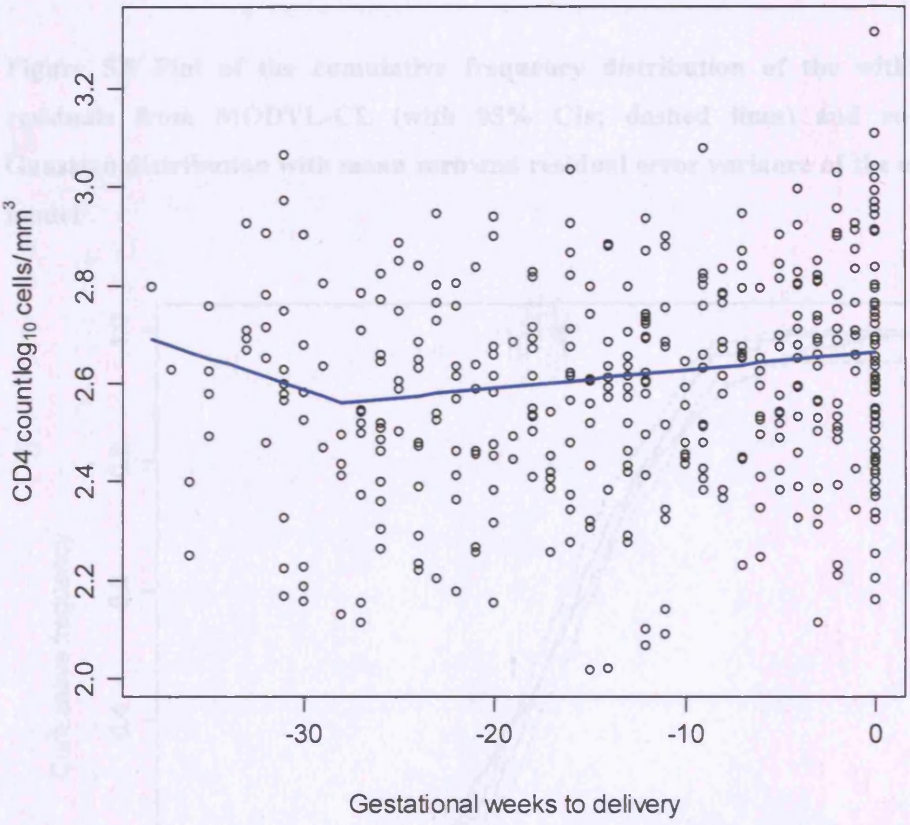
Allowing for a change in slope of CD4 count 28 weeks before delivery, CD4 count was estimated to decrease by 0.013 $\log_{10}$ cells/mm^3 per week in the period up to 28 weeks before delivery and thereafter to increase by 0.0037 $\log_{10}$ cells/mm^3 per week up to the time of delivery (Table 5.6). This equates to a 3.0% decrease in cells/mm^3 per week

(95%CI -0.69%, -4.5%) for the first period of pregnancy up to 28 weeks before delivery, followed by a 0.86% increase in cells/mm^3 per week (95%CI 0.46%, 1.16%) up to the time of delivery. Although the estimated change in CD4 count over pregnancy was small, both slopes were highly statistically significant (Table 5.6). Figure 5.7 shows the estimated slopes over pregnancy from MODCD4.

Black women had estimated CD4 counts 0.12 $\log_{10}$ lower than white women without a history of IDU, which equates to a 24% lower CD4 count (95%CI -11%, -55%), while levels were similar among white women irrespective of their history of IDU.

The significant correlation between the random effects for the slope and intercept was large and revealed that the higher the CD4 count at the change point (28 weeks before delivery), the smaller the increase in CD4 counts up to the time of delivery ($\hat{\rho} = -0.37$). The estimated standard deviation for the random effects for the slope and intercept were both bounded away from zero and the deviation for the slope was relatively small; this may lead to computational problems in the estimation procedure and in the computation of the integral for the bivariate random effects model fitted with the conditional likelihood approach (MODBIV-CL), especially with respect to the estimation of the covariance (correlation) between the random effects terms (Goldstein 1989).

Figure 5.7 Estimated mean change in log10 CD4 count



5.3.4 Model residuals from MODVL-CL and MODCD4

Both of the univariate LME models, MODCD4 and MODVL-CL, assume that the within-group residual errors are normally distributed with mean zero and constant variance (Pinheiro and Bates 2000;Jacqmin-Gadda *et al.* 2000). Figure 5.8 shows the Kaplan-Meier estimate of the cumulative frequency distribution of the computed right-censored residuals from MODVL-CL and the cumulative distribution of a normal variable with mean zero, and variance 0.66, i.e. the estimated within-group residual variance. The distribution of the residual error terms was similar to that of the assumed Gaussian distribution, but there were some visible departures in the tails of the two

distributions, where the curve of the assumed Gaussian distribution can be seen to fall just outside of the 95% CI for the within-group residuals.

Figure 5.8 Plot of the cumulative frequency distribution of the within-group residuals from MODVL-CL (with 95% CIs; dashed lines) and cumulative Gaussian distribution with mean zero and residual error variance of the estimated model

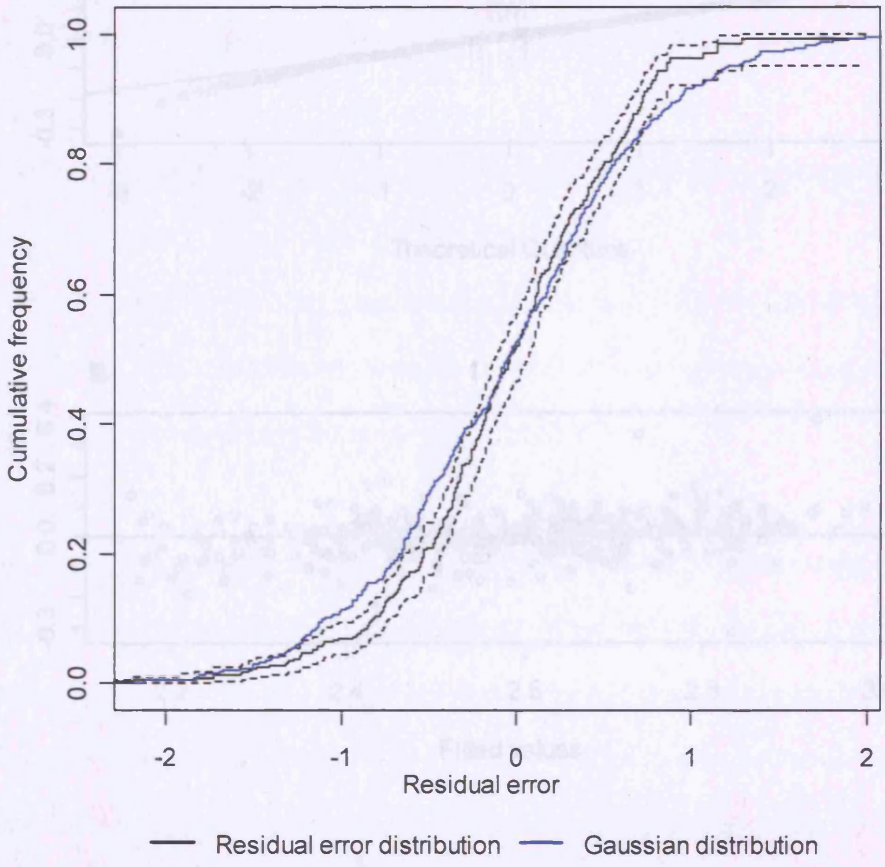
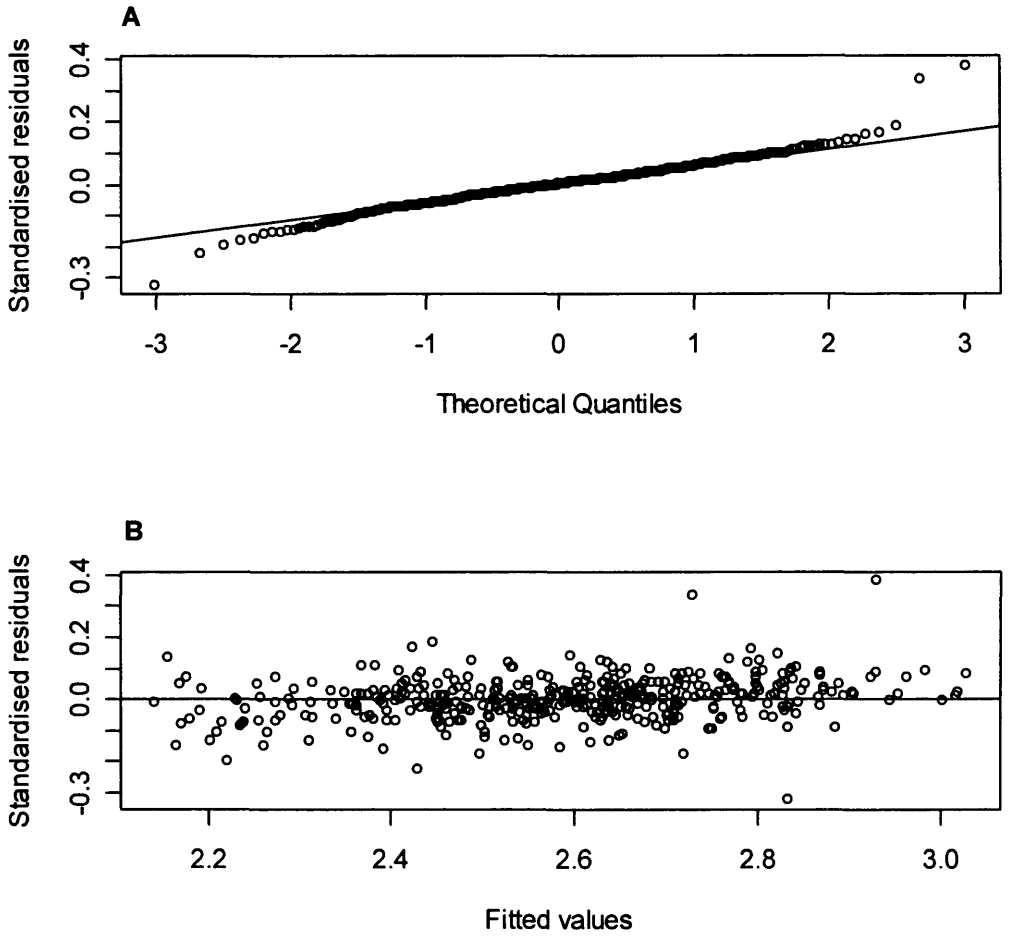


Figure 5.9 shows plots of the standardised within-group residuals from MODCD4

Figure 5.9 A) Normal probability plot of standardised residuals, B) Plot of standardised residuals against fitted values,MODCD4



The assumption of normality for the within-group errors was first assessed with a normal probability plot of the residuals, as shown in Figure 5.9A, where the assumed cumulative Gaussian distribution for these residuals is represented by the straight line. The distribution of the residuals follows the line fairly well, although there is evidence of some departure from normality in the tails. Figure 5.9B shows a plot of the within-

group errors against the fitted values; the residuals are mostly centred on zero with constant variance. However, the residuals for the tails appear to be centred slightly above or below zero, again suggesting a departure from normality caused by very few points in the tails of the distribution as seen in Figure 5.9A.

Overall, the assumption of a Gaussian distribution for the within-group residuals for the two univariate models appears plausible, which supports the accuracy of the estimated parameters and standard errors from the model and any inferences made.

5.3.5 Bivariate model for CD4 and viral load using midpoints for undetectable values (MODBIV)

MODBIV was fitted using the same covariates included in the two univariate models, MODVL and MODCD4. The estimated coefficients from the joint model, together with those from the univariate models are given in Table 5.7.

There were no substantial changes in the estimates from the bivariate model and their standard errors for the fixed effects, compared to those from the univariate model. The estimated parameters for the random effects were also unchanged, although there were some modest changes in the standard errors. This suggests that the information provided by the CD4 and viral load data contributes little to the estimation of the fixed effects estimates for either of the other markers in the joint model.

Table 5.7 Estimated coefficients from bivariate model and univariate models

	Bivariate model		Univariate model	
	Coefficient (95%CI)	***p***	**Coefficient (95%CI)**	***p***
	log_{10} viral load (MODBIV)		***log_{10} viral load (MODVL)***	
Mean at -28 weeks	4.12 (3.71, 4.53)	<0.001	4.12 (3.70, 4.53)	<0.001
Short term slope (<-28 weeks)	0.043 (-0.02, 0.11)	0.190	0.039 (-0.03, 0.10)	0.240
Long term slope (≥-28 weeks)	-0.055 (-0.06, -0.045)	<0.001	-0.055 (-0.06, -0.045)	<0.001
Race by IDU				
White non-IDU	0.00		0.00	
Black non-IDU	-0.20 (-0.47, 0.076)	0.160	-0.21 (-0.49, 0.06)	0.240
White IDU	-0.32 (-0.60, -0.04)	0.020	-0.31 (-0.59, -0.03)	0.030
Time period				
1998-1999	0.00		0.00	
2000-2001	-0.24 (-0.63, 0.15)	0.240	-0.24 (-0.64, 0.16)	0.250
2001-2006	-0.44 (-0.81, -0.07)	0.020	-0.44 (-0.82, -0.06)	0.025
Random effects – viral load				
Intercept	0.68 (0.52, 0.99)		0.69 (0.50, 0.93)	
Second slope	0.024 (0.015, 0.054)		0.024 (0.013, 0.042)	
Covariance between intercept and slope	-0.70 (-0.96, -0.44)		-0.70 (-0.88, -0.36)	

Table 5.7 contd.

	Bivariate model		Univariate model	
	Coefficient (95%CI)	***p***	**Coefficient (95%CI)**	***p***
	log_{10} CD4 (MODBIV)		***log_{10} CD4 (MODCD4)***	
Mean at -28 weeks	2.56 (2.51, 2.61)	<0.001	2.56 (2.51, 2.61)	<0.001
Short term slope (<-28 weeks)	-0.012 (-0.02, -0.003)	0.01	-0.013 (-0.02, -0.003)	0.008
Long term slope (≥-28 weeks)	0.0038 (0.002, 0.005)	<0.001	0.0037 (0.002, 0.005)	<0.001
Race by IDU				
White non-IDU	0.00		0.00	
Black non-IDU	-0.12 (-0.19, -0.05)	0.001	-0.12 (-0.19, -0.05)	0.001
White IDU	0.025 (-0.05, 0.09)	0.510	0.028 (-0.05, 0.10)	0.460
Random effects – CD4 counts				
Intercept	0.18 (0.16, 0.22)		0.18 (0.15, 0.22)	
Second slope	0.004 (0.003, 0.007)		0.004 (0.003, 0.006)	
Covariance between intercept and slope	-0.37 (-0.66, -0.09)		-0.37 (-0.62, -0.05)	

Table 5.8 shows the correlation matrix and 95% confidence intervals estimated from the bivariate model for CD4 count and viral load, using the Delta method (Lynch M. and Walsh B. 1998).

Table 5.8 Estimated correlations between viral load and CD4 (95% CI) from MODBIV

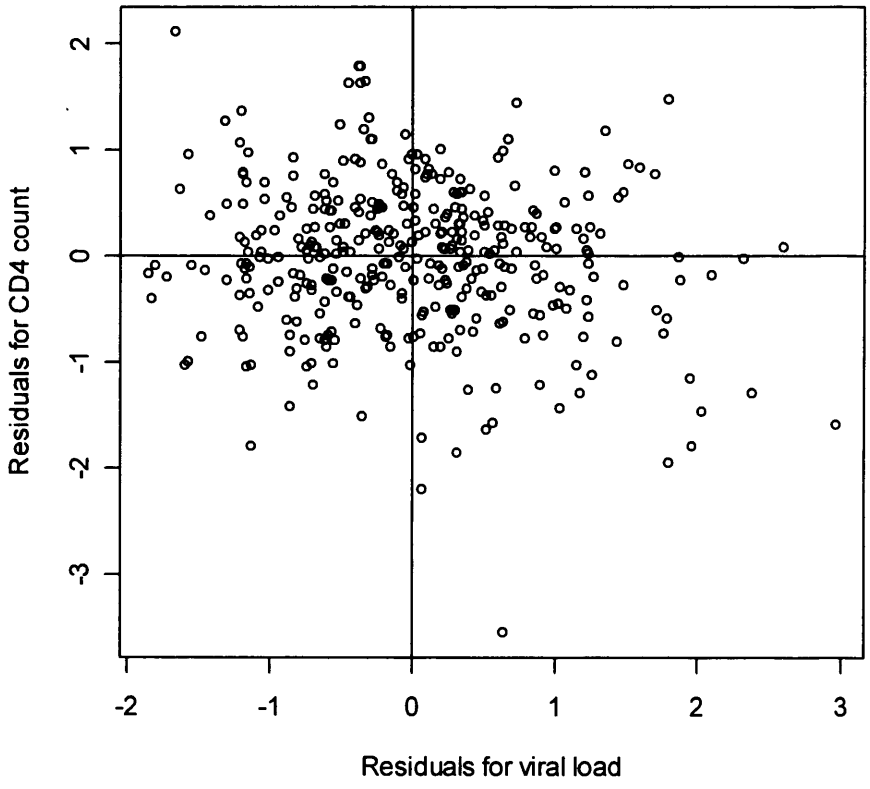
	Viral load intercept	CD4 count Intercept	Slope of viral load	Slope of CD4 count
Viral load intercept	1			
CD4 count intercept	-0.36 (-0.66, -0.07)	1		
Slope of viral load	-0.70 (-0.96, -0.44)	0.25 (-0.21, 0.71)	1	
Slope of CD4 count	0.27 (-0.21, 0.75)	-0.37 (-0.66, -0.09)	-0.29 (-0.99, 0.53)	1

There was evidence of correlation between the random intercept terms for the two markers, but not between any of the slopes (Table 5.8). The subject-specific intercepts for viral load and CD4 were negatively and significantly correlated $\left(\hat{\rho}_{CD4\,\mathrm{int}\,VL\,\mathrm{int}} = -0.36\right)$, i.e. the higher a woman's CD4 count at the time of change-point, the lower her mean viral load. The correlations between the other random effect terms were weak i.e. not statistically significant, and had wide confidence intervals. However, the estimates suggest the following possible correlations: a) the higher a woman's CD4 count 28 weeks before delivery, the more gradual the decrease in viral load to the time of delivery $\left(\hat{\rho}_{VLslopeCD4\,\mathrm{int}} = 0.25\right)$, b) the higher a woman's viral load 28 weeks before delivery, the steeper their increase in CD4 count to the time of delivery $\left(\hat{\rho}_{CD4slopeVL\,\mathrm{int}} = 0.27\right)$, and c) the steeper the decrease in the slope of viral load to the time of delivery, the steeper the increase in CD4 counts over the same period $\left(\hat{\rho}_{CD4slopeVLslope} = -0.29\right)$.

5.3.6 Bivariate residuals

Figure 5.10 shows a scatter plot of the bivariate residuals; the residuals were mostly scattered around zero, although there was some indication of a pattern in the lower right quadrant of the plot.

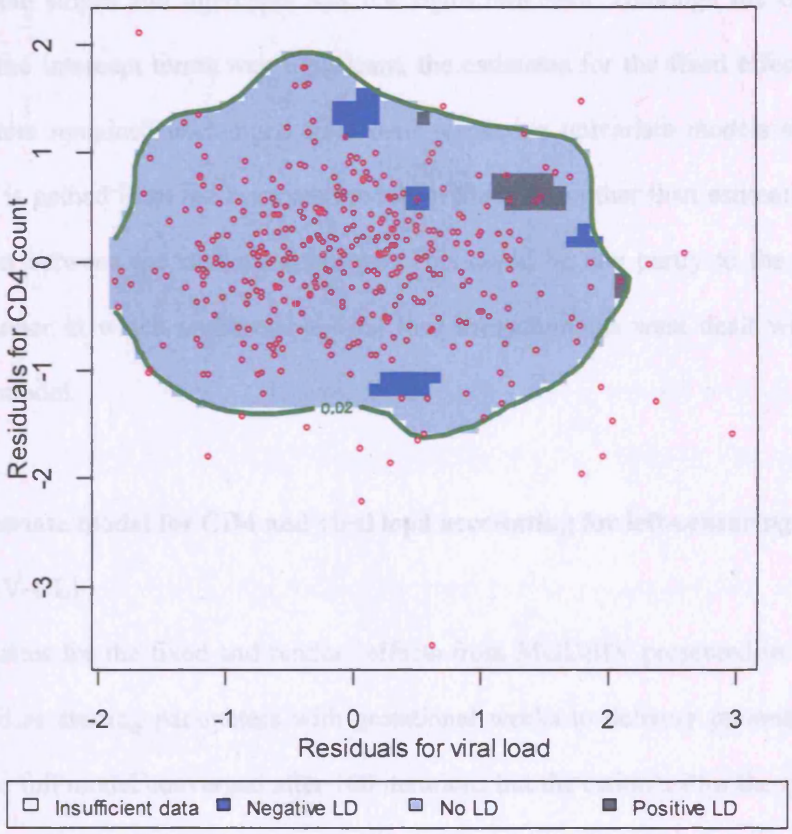
Figure 5.10 Scatter plot of the residuals from the bivariate model for CD4 and viral load



The corresponding Pearson correlation coefficient between the residuals was -0.13 (95%CI -0.23, -0.03). However, the Pearson correlation summarises the dependence between the residuals with a single scalar measure only and may be an unreliable summary measure over the whole range of residuals from the CD4 and viral load model.

The local dependence map (LDM) of the residuals was used to test for 'local' dependence between the residuals and is shown in Figure 5.11.

Figure 5.11 Local dependence map of the residuals from MODBIV



The LDM carries out localised Pearson correlation tests and provides an associated significance test for the local dependence between residuals, with the colours given in the figure legend indicating the presence of positive, negative or no dependence. Overall, the LDM revealed a very weak indication of local dependence between the residuals for viral load and CD4 counts, confirming the assumption of independence between the residuals from MODBIV.

As mentioned in section 5.2, the BLME model (MODBIV) accounts for the correlation between the two markers through the covariance of their random effects terms; however, the estimated correlation between the random effects terms for the slopes, or between the slopes and intercepts was not significant here. Although the correlation between the intercept terms was significant, the estimates for the fixed effects for the two markers remained unchanged from their respective univariate models suggesting that little is gained from the bivariate model for these data other than estimation of the correlation between the random intercepts. This could be due partly to the relatively crude manner in which undetectable viral load measurements were dealt with in this bivariate model.

5.3.7 Bivariate model for CD4 and viral load accounting for left-censoring (MODBIV-CL)

The estimates for the fixed and random effects from MODBIV presented in Table 5.7 were used as starting parameters with gestational weeks to delivery parameterised in years. The full model converged after 160 iterations but the estimates for the covariance elements of the random effect terms (corresponding to the correlation coefficients) were very close to the starting parameters and none of their standard errors had been meaningfully estimated.

The Quasi-Newton optimisation procedure used for this model is a gradient method, where the final estimated parameters are such that they maximise the partial-derivatives of the likelihood function with respect to each parameters, i.e. corresponding to the point where the gradient of the likelihood function is zero (Everitt 1987). However,

although the SAS procedure converged successfully, the final model resulted in elements of this final gradient to be >|0.001| for 12 of the 24 estimated parameters with some gradients ranging from as high as 0.05 to 0.20. This indicated that the estimated parameters were not those that optimise the likelihood suitably, and may be far from the true optimal values (SAS Institute Inc. 2004). Additionally, the final estimated Hessian matrix was not positive definite; this occurs when the likelihood is maximised by setting one or more of the variance components to zero (SAS Institute Inc. 2004). As a consequence, the inverse of the Hessian matrix, i.e. the estimated variance-covariance matrix, may be unreliable for some or all of the fixed and random parameter estimates.

The estimates of the correlations between the slope and intercept terms for the two markers were small and not significant in MODBIV (Table 5.8). Goldstein suggests omitting these poorly estimated parameters to stabilise the estimates for the remaining parameters, particularly for estimation of the random parameters when the sample size is relatively small (Goldstein 1989). Therefore, MODBIV-CL was refitted with the same covariates, but with only the random intercepts for the two markers and the covariance between them. This model converged, but still resulted in gradients >|0.001| for some of the parameters and a Hessian matrix which was not positive-definite. Attempts to fit the model with gestational age only or with random effects for the slopes only also resulted in unstable estimates.

One of the principal reasons for using the bivariate model here was to improve the estimates for the fixed effects for both viral load and CD4 through estimation of the correlation between them, while also accounting for the presence of left-censoring. An

alternative approach, modelling viral load and CD4 count in univariate LME models as a function of the other marker, was considered.

5.3.8 Univariate viral load model accounting for left-censoring as a function of CD4 count (MODVL-CL2)

This analysis included the 143 women from the previous models. There were 367 measurements available on viral load and CD4 count taken at the same time in pregnancy, with an additional 20 measurements where a CD4 count was taken within 4 weeks of the viral load measurement; 43 viral load measurements without a close CD4 measurement were excluded. The median number of measurements per woman was 3 (IQR 2-3).

MODVL-CL (Table 5.5) was fitted again with the same random effects (for the intercept and slope) and the fixed effects, with the addition of $\log_{10}$ CD4 count and an interaction between $\log_{10}$ CD4 and both slopes. The removal of the interaction term between the second slope and CD4 did not result in a significant increase in the log-likelihood $\left(\text{LRT: } \chi^2_{df=1} = 0.5; p = 0.48\right)$, hence it was excluded. Although the Wald tests for the covariate for time period was not significant at the 5% level, their removal led to a significant worsening of the fit of the model and were therefore retained $\left(\text{LRT: } \chi^2_{df=1} = 7.6; p = 0.02\right)$. The fixed and random effects estimates from the final model are given in Table 5.9.

Table 5.9 Adjusted coefficients from MODVL-CL2 (n=143)

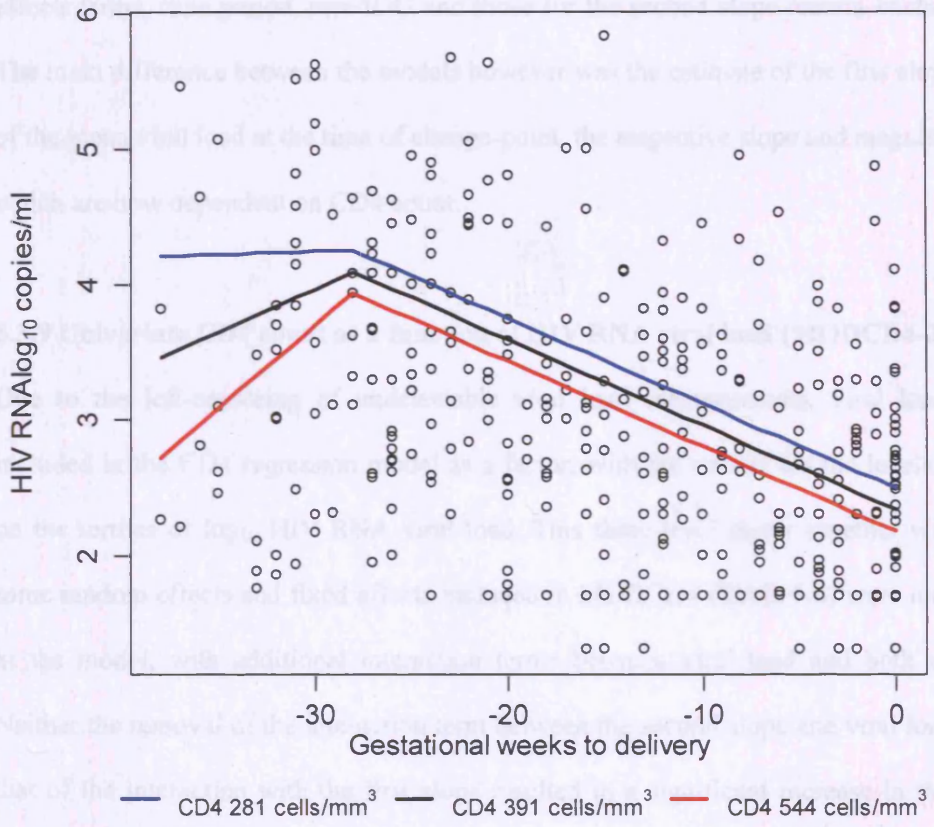
	$\log_{10}$ viral load Coefficient (95%CI)	p
Initial slope (<-28 weeks)	-1.00 (-1.73, -0.27)	<0.001
Slope to delivery (>-28 weeks)	-0.063 (-0.075, -0.05)	<0.001
$\log_{10}$ increase in CD4	-1.11 (-1.77, -0.45)	0.001
$\log_{10}$ increase in CD4*initial slope	0.41 (0.13, 0.69)	0.005
Race by IDU		
White non-IDU	0.00	
Black non-IDU	-0.34 (-0.68, 0.009)	0.06
White IDU	-0.38 (-0.72, -0.04)	0.03
Time period		
1998-1999	0.00	
2000-2001	0.043 (-0.44, 0.53)	0.86
2002-2006	-0.36 (-0.82, 0.11)	0.13
Random effects parameters		
Intercept	0.65 (0.33, 0.86)	
Slope to delivery	0.033 (0.003, 0.046)	
Correlation between intercept and slope	-0.41 (-0.85, 0.03)	
Within-group residual deviation	0.75 (0.64, 0.84)	

The overall level of viral load was dependent on CD4 count in the adjusted model (an estimated mean HIV RNA level 1.11 $\log_{10}$ copies/ml lower for each unit increase in $\log_{10}$ CD4 cells/mm^3). The change in viral load over the period up to 28 weeks before delivery was also dependent on CD4 count (Table 5.9), and thereafter was estimated to

decrease by 0.063 $\log_{10}$ copies/ml per week (95%CI -0.075, -0.05) until the time of delivery independent of CD4 count. The change in viral load after 28 weeks to the time of delivery corresponds to a 14% weekly decrease in copies/ml (95%CI -11%, -16%) until the time of delivery. Due to its dependence on CD4 count, the change in the initial slope in pregnancy is more easily interpreted graphically. Figure 5.12 shows the estimated slopes over pregnancy by CD4 count at 28 weeks before delivery; the blue line represents women with CD4 counts of 281 cells/mm^3 (lower quartile), the black line represents women with CD4 counts of 391 cells/mm^3 (median) and the red line, women with CD4 counts of 544 cells/mm^3 (upper quartile). This figure depicts the relationship between the initial slope of viral load in pregnancy and CD4 count; women with the highest CD4 count were estimated to have the steepest increase in the initial period, with women with the lowest CD4 count experiencing little change in viral load over the same period. It should be noted here that although the 95% CIs for the initial slope and the interactions between CD4 and this slope did not contain zero, they were relatively large, reflecting the relatively small number of measurements contributing to the estimation of the initial slope.

White women with a history of IDU were estimated to have a 0.38 $\log_{10}$ lower viral load than white women with no history of IDU (Table 5.9); this equates to a 58% lower viral load (95%CI -8.8% -81%). Black women were also estimated to have a similar magnitude of viral load below that of white women with no history of IDU, but this was only borderline significant.

Figure 5.12 Estimated mean change in log_{10} viral load by CD4 count from MODVL-CL2



There estimated random effects terms reveal the considerable variation among the intercept and slopes between women; the estimated correlation between the two random terms was moderate, but the associated confidence interval was bounded by zero. This suggests that the estimated slope of viral load to delivery is relatively invariant to deviations from the estimated mean; i.e. women with mean viral loads above or below the estimated intercept will experience similar changes in viral load over the 28 weeks to the time of delivery.

It is not appropriate to make direct comparisons of the estimates with the univariate model which did not adjust for CD4 counts (Table 5.5) as the two models are based on slightly different datasets, but it is interesting to see that the estimates for the random effects terms, time period, race/IDU and those for the second slope remain unchanged. The main difference between the models however was the estimate of the first slope and of the mean viral load at the time of change-point, the respective slope and magnitude of which are now dependent on CD4 count.

5.3.9 Univariate CD4 count as a function of HIV RNA viral load (MODCD4-2)

Due to the left-censoring of undetectable viral load measurements, viral load was included in the CD4 regression model as a factor, with the cut-off for the levels based on the tertiles of $\log_{10}$ HIV RNA viral load. This three-level factor together with the same random effects and fixed effects included in MODCD4 (Table 5.6) were included in the model, with additional interaction terms between viral load and both slopes. Neither the removal of the interaction term between the second slope and viral load, nor that of the interaction with the first slope resulted in a significant increase in the log-likelihood $\left(\text{LRTs: } \chi^2_{df=2} = 2.44; p = 0.29, \text{ and } \chi^2_{df=2} = 0.42; p = 0.81\right)$, hence both terms were excluded. The fixed and random effects estimates from the final model are given in Table 5.10

Table 5.10 Adjusted coefficients from MODCD4-2 (n=143)

	$\log_{10}$ CD4 count Coefficient (95%CI)	p
Mean at 28 weeks before delivery	2.64 (2.58, 2.70)	<0.001
Initial slope (<-28 weeks)	-0.007 (-0.015, 0.0017)	0.120
Slope to delivery (>-28 weeks)	0.0014 (-0.0001, 0.0029)	0.077
Tertiles of HIV RNA		
< 2.30 $\log_{10}$ copies/ml	0.00	
2.30-3.25 $\log_{10}$ copies/ml	-0.026 (-0.056, 0.0043)	0.100
≥3.25 $\log_{10}$ copies/ml	-0.072 (-0.11, -0.035)	<0.001
Race by IDU		
White non-IDU	0.00	
Black non-IDU	-0.13 (-0.21, -0.062)	<0.001
White IDU	0.021 (-0.053, 0.094)	0.580
Random effects parameters		
Intercept	0.19 (0.17, 0.22)	
Slope to delivery	0.0046 (0.0034, 0.0063)	
Correlation between intercept and slope	-0.45 (-0.65, -0.20)	
Within-group residual deviation	0.079 (0.070, 0.089)	

Allowing for viral load and a change in slope of CD4 count 28 weeks before delivery, CD4 count was estimated to be relatively stable up to 28 weeks before delivery (Table 5.10). CD4 counts were estimated to increase by 0.0014 $\log_{10}$ cells/mm^3 per week in the subsequent period; this equates to a 0.32% weekly increase in cells/mm^3 (95%CI -0.023%, 0.67%) up to the time of delivery, but was of borderline significance. Although

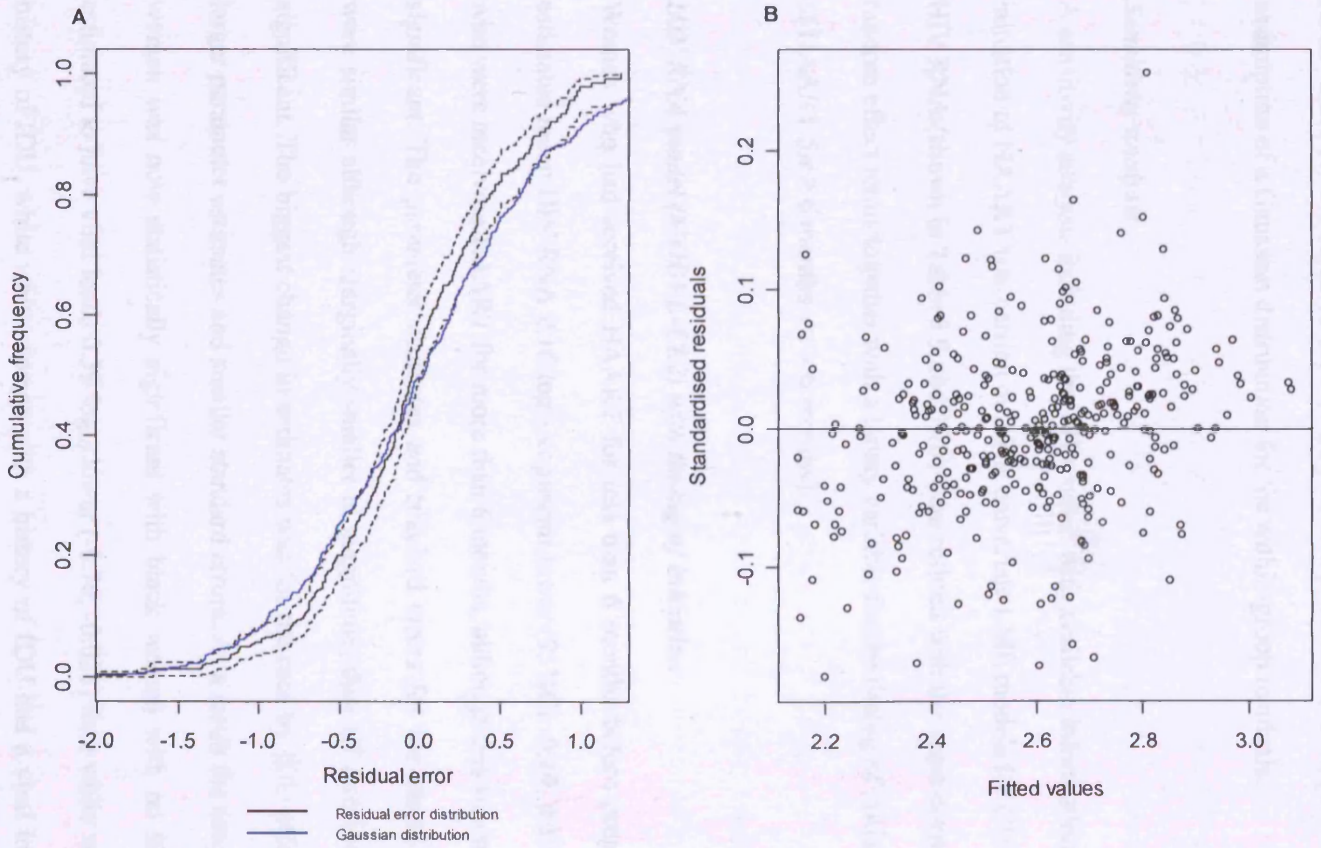
there was no evidence of any dependence between the slope of CD4 over pregnancy and viral load, the mean level of CD4 count was 0.072 $\log_{10}$ lower with viral loads above 3.25 $\log_{10}$ copies/ml, which equates to a mean CD4 count 15% lower (95%CI -7.7%, -22%). That is, women with the highest viral load had consistently lower CD4 counts than women with viral loads below 2.30 $\log_{10}$ copies/ml, over the period of pregnancy.

The association between lower CD4 counts and black race was also significant, with this group of women estimated to have CD4 counts which were 0.13 $\log_{10}$ or 26% lower than white women without a history of IDU (95%CI -13%, -38%), while levels were similar among white women irrespective of their history of IDU.

Adjusting for viral load in MODCD4-2 resulted in CD4 slopes which were no longer significant, compared to those in MODCD4 (Tables 5.6 and 5.10). There were only small changes in the estimates for the remaining fixed effects terms and random effects terms. However, the estimated correlation between the two random effects terms remained negative and statistically significant but was considerably stronger here than in MODCD4. The correlation suggests that the higher the CD4 count for a woman at the intercept, the more gradual the subsequent increase in CD4 over pregnancy (Table 5.10).

The Kaplan-Meier estimate of the cumulative frequency distribution of the computed right-censored residuals from the adjusted viral load model are shown in Figure 5.13A.

Figure 5.13 Residuals from univariate models for: A) Viral load adjusted for CD4 count and accounting for left-censoring, B) CD4 count adjusted for viral load level



The within-group residuals here appear to satisfy the assumption of normality better than MODVL-CL, with the Gaussian curve sitting closer inside the 95% CIs in the tails of the distribution. Figure 5.13B shows a plot of the within-group errors against the fitted values for the adjusted CD4 model; the two plots reveal no major violations in the assumption of a Gaussian distribution for the within-group residuals.

Sensitivity analysis

A sensitivity analysis including the 109 women with available information on timing of initiation of HAART was carried out. The univariate LME models for CD4 count and HIV RNA (shown in Table 5.9 and 5.10) were refitted with the same covariates and random effect terms together with a binary variable for the timing of initiation (receipt of HAART for ≥ 6 months vs. < 6 months).

HIV RNA model (MODVL-CL2) with timing of initiation

Women who had received HAART for less than 6 months before pregnancy had an estimated mean HIV RNA 0.16 $\log_{10}$ copies/ml lower (95%CI -0.49, 0.17) than women who were receiving HAART for more than 6 months, although this was not statistically significant. The parameter estimates and standard errors for the slopes in pregnancy were similar although marginally smaller in magnitude; they all remained statistically significant. The biggest change in estimates was for the race by IDU effect, which had larger parameter estimates and smaller standard errors. As a result the estimate for black women was now statistically significant with black women with no history of IDU estimated to have viral loads 0.39 $\log_{10}$ lower (-0.76, -0.023) than white women with no history of IDU, while white women with a history of IDU had a viral load which was 0.52 $\log_{10}$ lower (95%CI-0.89, -0.15).

CD4 model (MODCD4-2) with timing of initiation

Women who had received HAART for less than 6 months before pregnancy had an estimated mean CD4 0.040 $\log_{10}$ cells/mm^3 higher (95%CI -0.029, 0.11) than women who were receiving HAART for more than 6 months, although this was not statistically significant. The parameter estimates and standard errors for the initial slope in pregnancy, race and viral load categories were similar. However, the slope from -28 weeks up to the time of delivery was now marginally larger and the associated standard error was smaller (0.0017 $\log_{10}$ cells/mm^3 per week increase, 95%CI 0.0006, 0.0034).

In conclusion, the bivariate model (MODBIV) was able to estimate the relationship between the two markers over the whole period of pregnancy, without assuming any dependence of one marker in relation to the other. However, it had little impact on the fixed effects estimates. This suggests that accounting for the relationship between the two markers through their random effects covariance may not be sufficient to explain important variation in the markers resulting from their joint distribution. A bivariate model with left-censoring could not be estimated and therefore adjustment for the two markers within a univariate model for CD4 count and HIV RNA viral load accounting for left-censoring (MODCD4-2 and MODVL-CL2) was used instead which resulted in marked differences in the estimates of the fixed-effects. For this data therefore, MODCD4-2 and MODVL-CL2 were the preferred models.

5.4 Key points

- Mean plots of CD4 and HIV viral load suggested a two-slope evolution of the markers over pregnancy.
- The change in slope for both markers was estimated to occur around the end of the first trimester of delivery, around 28 weeks before the time of delivery.

Univariate HIV RNA viral load model independent of CD4 counts (MODVL-CL)

- The slope of viral load over pregnancy was underestimated in the model which did not account for left-censoring appropriately.
- When CD4 was not adjusted for, viral load was estimated to be relatively stable over the beginning of pregnancy, thereafter decreasing by a weekly 15% until the time of delivery.
- White women with a history of IDU were estimated to have lower viral loads than white women with no history of IDU.
- The estimated correlation between the random terms suggested that the estimated slope of viral load to the time of delivery was relatively invariant to deviations from the estimated mean.

Univariate HIV RNA viral load model as a function of CD4 counts (MODVL-CL2)

- The initial slope of viral load in pregnancy was found to be dependent on a woman's CD4 count; women with low CD4 count had relatively stable changes in viral load over this period while women with high CD4 counts experienced a steeper increase in this period.

- The level of CD4 count was associated with the mean viral load resulting in shifts in the slope of viral load, but was found to be independent of the second slope itself.
- There were little changes in the estimates of viral load for the second slope and for race between the two models.

Univariate CD4 model independent of HIV RNA viral load (MODCD4)

- CD4 count was estimated to decrease in the first period of pregnancy up to 28 weeks before the time of delivery, and thereby to increase by 0.9% each week until the time of delivery.
- Black women were estimated to have CD4 counts which were 24% lower than white women.

Univariate CD4 model as a function of HIV RNA viral load (MODCD4-2)

- Adjusting for viral load, CD4 count was estimated to be stable up to 28 weeks before delivery, and to then increase by 0.32% each week up to the time of delivery, although this change was only borderline significant.
- Large viral loads (greater than 3.25 $\log_{10}$ copies/ml) were associated with significant decreases in the mean CD4 count but there was no association between viral load and either of the slopes during pregnancy.
- The estimated correlation between the random terms was strong and suggested that higher the CD4 count for woman 28 weeks before delivery, the more gradual their subsequent increase in CD4 over pregnancy.

Bivariate model for HIV RNA viral load and CD4 count (MODBIV)

- The bivariate model assuming midpoints for undetectable viral load measurements resulted in negligible changes in the estimates and standard errors of the fixed and random effects terms from those from the univariate models.
- There was evidence of substantial negative correlation between the intercepts for the two markers, such that the higher a woman's CD4 count at the time of change-point, the lower their mean viral load.
- The correlations between the two slopes, or between the slopes and intercepts were weak and not significantly different from zero.
- Attempts to fit a bivariate model accounting for left-censoring were unsuccessful, resulting in unstable estimates of the model parameters and associated standard errors.

Chapter 6 ART use during pregnancy and adverse neonatal outcomes

6.1 ART use during pregnancy and risk of congenital abnormalities

6.1.1 Introduction

The use of HAART during pregnancy, particularly in the early stages when organogenesis is taking place, has raised concerns relating to the safety for the exposed fetus (Public Health Service Taskforce 2006b). However, few data are available to ascertain whether the risk of congenital abnormalities is increased by first trimester exposure to antiretrovirals and/or by certain combinations of treatment, including HAART.

The aim of this analysis was to assess the possible teratogenic risks related to in utero exposure to any antenatal ART, by investigating the prevalence of congenital abnormalities among infants born to HIV-infected women by earliest exposure to and by type of treatment regimen among infants.

6.1.2 Methods specific to this analysis

Data

This analysis included data on 3740 mother-child pairs enrolled in the ECS by December 2003.

To assess the outcome of pregnancies by type of treatment regimen, ART exposures were grouped into four categories, NRTI only, combinations of NRTI and PI, NRTI and NNRTI or a combination of NRTI, NNRTI and PI. As organogenesis occurs in the first trimester, rates of congenital abnormalities among infants with second/third trimester

exposures to ART were grouped together and compared with those with first trimester exposures. The first trimester of pregnancy was defined as up to and including 12 gestational weeks.

Infants were assessed for the presence of congenital abnormalities at birth by a senior obstetrician and/or a senior paediatrician; additional perinatal assessments were made by experienced paediatricians, who were also responsible for reports on further paediatric follow up visits.

The association between type of antenatal ART and the presence of congenital abnormalities was assessed with a logistic regression model, after adjusting for information available on IDU in pregnancy (Y/N) and maternal age at delivery (<25, 25-34 and ≥35 years).

The analysis was carried out using R version 1.9.0 (R Development Core Team 2006).

6.1.3 Results

Of the 3740 infants enrolled at the time of this analysis, 1973 (53%) were exposed to ART in utero, including 602 to HAART. Of these 1973 infants, 789 (40%) were conceived while their mother was taking ART and exposed in the first trimester, and the remainder exposed from the second and third trimester (n=702 and n=482, respectively). Median maternal age at delivery was 28 years (range: 10-45 years). Most mothers were white (72%, 2700/3740), 21% (781/3740) were black (mainly from sub-Saharan Africa) and 7% (256/3740) were from other ethnic groups. Thirteen percent (493/3740) of infants were exposed to illicit drugs during pregnancy. CD4 counts were

available for 2042 (55%) mothers; median CD4 count at delivery was 420 cells/mm^3 (range: 280-600 cells/mm^3). Median gestational age at delivery was 38 weeks (range 22– 43 weeks) and median birth weight was 2940g (range 500 – 5190g).

Congenital abnormalities

Congenital abnormalities were recorded in 1.5% (55/3740) of children, 31 (56%) of whom had been exposed to ART in utero, including 14 in the first trimester of pregnancy. Table 6.1 gives a summary of the congenital abnormalities by exposure group. Among the 31 exposed children, abnormalities included ear malformations (2), cleft palate (2), ventricular septal defect (6), atrial septal defect (2), blood-clotting disorder (1), hydrocele in both testes (1), polydactyly (3), Down's syndrome (3), polycystic kidney (3), hydronephrosis (2), oesophageal atresia (2), ileostoma and enteritis (1), gastric perforation (1), hydrocephalus (1 died four days after birth), cataract (1) and situs inversus, transposition of arteries (1). Congenital abnormalities in the 24 unexposed children included cleft palate (1), ventricular septal defect (4), cardiac rhabdomyoma (1), tetralogy of fallot (1), polydactyly (2), syndactyly(1), polycystic kidney (2), hydronephrosis (2), oesophageal atresia (3), microcephalus (1), hydrocephalus (1), hip dysplasia (2), cataract (1), situs inversus, transposition of arteries (1), and diaphragmatic hernia (1).

Table 6.1 Summary of congenital abnormalities by exposure to ART

	No Antenatal exposure To ART	Earliest Antiretroviral Therapy (ART) Exposure in First Trimester					Earliest ART Exposure 2nd and/or 3rd Trimester
		NRTI(s) Only	PI(s) + NRTI(s)	NRTI(s) + NNRTI(s)	PI(s) + NRTI(s) + NNRTI(s)	Overall First Trimester Exposure	
Mother-child pairs (n)	1767	280	273	195	41	789	1,184
Children with congenital abnormalities (n)	24	1	7	6	0	14	17
Number of abnormalities:							
EAR, FACE AND NECK	0	0	0	0	0	0	2
CLEFT PALATE	1	0	0	0	0	0	2
VENTRICULAR SEPTAL DEFECT	4	1	1	1	0	3	3
OTHER HEART DEFECTS	2	0	1	1 †	0	2	0
OTHER CIRCULATORY SYSTEM	0	0	1	0	0	1	0
MALE GENITALIA	0	0	0	1 †	0	1	0
LIMB REDUCTION/ADDITION	3	0	0	0	0	0	3
DOWN'S SYNDROME	0	0	0	0	0	0	3
RENAL AND KIDNEY	4	0	1	2	0	3	2
GASTROINTESTINAL	3	0	2	2	0	4	0
NEUROLOGICAL	2	0	0	0	0	0	1
HIP DYSPLASIA	2	0	0	0	0	0	0
CATARACT	1	0	0	0	0	0	1
OTHER	2	0	1	0	0	1	0

† - one child with two defects; hydrocele in both testes and atrial septal defect

Type and timing of ART during pregnancy

There was no apparent increase in any particular abnormality with the use of ART or HAART during pregnancy. Among the 3740 infants, 1.4% of those not exposed to antenatal ART (24/1767) had a congenital abnormality versus 1.6% of those exposed (31/1973; χ^2=0.16, p=0.69). There was also no discernible pattern of congenital abnormalities between the type and timing of ART during pregnancy (Table 6.1).

The prevalence of congenital abnormalities in children exposed to any antenatal ART in the first trimester of pregnancy (14/789 [1.8%, 95%CI 0.97-3.0]) was similar to those exposed in the second or third trimester (17/1184 [1.4%, 95% CI 0.84 – 2.3]) (χ^2=0.17, p=0.68); the relative risk for first trimester exposures compared to later exposures was 1.3 (95%CI 0.6-2.5). Among the 789 infants with first trimester exposure, there was no difference in the prevalence of congenital abnormalities among infants exposed to HAART (11/546, 2.0% [95%CI 1.0-3.6]) compared with those exposed to mono/dual therapy only (3/243, 1.2% [95%CI 0.26-3.6]) (Fisher's exact test p=0.57).

In multivariable logistic regression analysis involving 3471 mother-child pairs with information on all variables, the presence of congenital abnormalities was not associated with the use of mono/dual therapy or HAART during pregnancy, after adjusting for maternal age at delivery and IDU during pregnancy (adjusted odds ratio (AOR) 1.05, [95%CI 0.49, 2.2, p=0.91] and 1.22, [95%CI 0.57, 2.6, p=0.61] for mono/dual therapy and HAART respectively with 'no antenatal ART' as the reference category).

In a sub-analysis among 1810 mother-child pairs with antenatal ART use, after adjusting for maternal age at delivery, there was no increased risk of congenital abnormalities associated with first trimester exposure to ART compared with later exposures (AOR

1.33, [95%CI 0.64, 2.8, p=0.44]). None of the mothers of children with congenital abnormalities used injecting drugs during pregnancy, precluding adjustment for this variable here.

Nineteen women in this cohort received efavirenz-containing HAART regimens at the time they became pregnant and continued taking the drug for a median of 40 days into their pregnancy (range 24-106 days). No congenital abnormalities (0%, 95% CI 0, 17.6) were reported in this group, but with the small numbers in this exposure group an increased risk cannot be excluded.

6.2 ART use during pregnancy and the risk of premature delivery

6.2.1 Introduction

European studies have confirmed the association between antenatal HAART use and premature delivery (ECS and Swiss Cohort and European Collaborative Study and the Swiss HIV Pregnancy Cohort 2000;European Collaborative Study 2004;Townsend *et al.* 2006). Data from a prospective study in America also reported an increased risk of premature delivery with the use of PI-based HAART, compared with HAART without (Cotter *et al.* 2006). Recent studies from Brazil and the US have also examined the impact of PI-based HAART on increased risk of prematurity, but with respect to none or mono/dual therapy only (Szyld *et al.* 2006;Schulte *et al.* 2007). However, data are lacking on whether the use of PI-based HAART increases the risk of prematurity over and above that associated with NNRTI-based HAART.

The aim of this analysis was to assess whether the use of antenatal PI-based HAART regimens is associated with an increased risk of premature delivery compared to NNRTI-based HAART in a European population of HIV-infected women.

6.2.2 Methods specific to this analysis

Data

This analysis was restricted to women who were receiving HAART at conception or who received HAART during pregnancy and who delivered a live born singleton infant with known gestational age. HAART was defined as a regimen of ≥3 ARV drugs, consisting of at least 2 NRTIs and at least one PI or NNRTI; women receiving both NNRTIs and PIs in pregnancy were excluded.

By June 2006, 1493 mother-child pairs enrolled in the ECS met these criteria. Ten were stillborn, 26 were missing gestational age at delivery and an additional 42 women received HAART with both NNRTI and PI and were excluded from the analysis. A further 121 were of a multiple birth and were excluded due to the higher preterm birth rate among multiple pregnancies (Slattery and Morrison 2002). A total of 199 mother-child pairs were therefore removed from the analysis and the study population for this analysis consisted of 1294 pairs.

Mode of delivery was categorised as vaginal or caesarean section (CS) delivery, with CS before labour and rupture of membranes (ROM) defined as elective CS and after labour/ROM as emergency CS.

Type of HAART regimen (NNRTI-containing or PI-containing) was considered in the regression model along with the following variables, which have elsewhere been shown to be important predictors for preterm delivery (ECS and Swiss Cohort and European Collaborative Study and the Swiss HIV Pregnancy Cohort 2000;European Collaborative Study 2004;Cotter *et al.* 2006;Patel *et al.* 2004;Slattery and Morrison 2002); timing of HAART initiation (before pregnancy, first, second or third trimester), race (white, black or other), CD4 cell count closest to the time of delivery (<200, 200-499, ≥500 cells/mm^3), maternal age (≤25, 26-34, ≥35 years), IDU during pregnancy (Y/N) and parity (0, 1, 2, ≥3 previous live births).

Statistical methods

European guidelines for management of pregnant HIV-infected women recommend elective CS deliveries, to be scheduled at 38 weeks, for PMTCT and the rate of elective CS in the ECS has therefore been high in recent years (65.4% in the period 2000-2004)

(European Collaborative Study 2004). However, some centres in the ECS had a policy of scheduling elective CS at 36 weeks (Newell *et al.* 2002;Brockmeyer 1999). This large number of elective CS deliveries, some carried out before 37 weeks, complicates the investigation of the association between antenatal HAART use and premature delivery. The use of χ^2 tests and logistic regression models to assess differences between treatment groups with respect to premature delivery could lead to spurious results, as these tests would not distinguish between significant differences due to prematurity or early elective CS. Removing the elective CS deliveries would allow for comparisons of prematurity between treatment groups for "natural deliveries" only (i.e. those excluding elective CS), but would exclude a large number of mother-child pairs from the analysis and could bias any inferences made from the model. Deliveries which start with the onset of labour and/or rupture of membranes can therefore be considered as spontaneous deliveries, regardless of whether the delivery was vaginal or emergency CS, while an elective CS delivery can be considered as an 'unspontaneous' delivery (whether scheduled before or after 37 weeks) as in principle the infant could have had a longer gestation than that observed.

A different approach is to examine the association of HAART exposure and prematurity by comparing gestational ages at delivery between groups, while accounting for the right-censoring of elective CS deliveries. This is more appropriate than the approach outlined above as it makes use of all available information. Therefore a survival analysis approach was taken to assess the association between type of antenatal HAART exposure and the hazard of delivery at different gestational ages, with deliveries by vaginal and by emergency CS taken to be observed events and those by elective CS as censored events. This is based on the risk sets of women who are still being followed up

at each gestational age at which a woman delivers and therefore accounts for the censored deliveries as opposed to simply removing them.

Survival curves obtained with the Kaplan-Meier method (Kaplan and Meier 1958) were used to explore the proportion of women delivering at different gestations, stratified by HAART group. Differences between survival curves were tested with a log-rank test (Therneau and Grambsch 2000). A Cox proportional-hazards model was used to estimate unadjusted and adjusted relative hazards (RH and ARH, respectively) of delivery by HAART group, after adjusting for the potential confounders mentioned above (Cox DR and Oakes D 1984;Therneau and Grambsch 2000). The assumption of proportional hazards both by the individual covariates, and for the model as a whole, was formally assessed with tests based on the examination of scaled Schoenfeld residuals over time (Therneau and Grambsch 2000;Grambsch and Therneau 1994).

The analysis was carried out using R version 2.3.1 (R Development Core Team 2006).

6.2.3 Results

The majority of the 1294 women (760/1294; 59%) received a PI-based HAART regimen, of whom 479 (65%) received nelfinavir (NFV). Of the 534 women receiving a NNRTI-based HAART regimen, 94% received nevirapine (NVP) as their NNRTI component. The characteristics of the mother-child pairs by type of HAART exposure are given in Table 6.2. There were no statistically significant differences between treatment groups with respect to most characteristics. However, the proportion of women delivering before 2001 was higher for the PI group and there were some differences in the proportion of women beginning treatment in the second and third trimester between treatment groups.

Prematurity

The overall median gestational age was 38 weeks (range 23-42), the prevalence of premature and severlye premature deliveries was 21.9% (284/1294, 95%CI 20-24%) and 5.6% (72/1294, 95%CI 4-7%) respectively and was similar between the two HAART groups when tested with a χ^2 test (Table 6.2).

Figure 6.1 shows a cumulative plot of the estimated proportion of women delivering at different gestational ages by Kaplan Meier analysis according to type of HAART exposure. Overall, the plot revealed little evidence of differences in the proportion of women delivering by exposure group, especially for gestations $\leq$37 weeks.

Table 6.2 Characteristics by HAART category

Characteristic	NNRTI-based HAART† (*n*= 534)	PI-based HAART† (*n* = 760)	*p* value‡
Maternal age			
Median (IQR)	31 (27-35)	31 (27-35)	0.38
≤ 25 years	99 (19)	123 (17)	
26 – 34 years	286 (54)	391 (54)	
≥35 years	141 (23)	214 (29)	0.50
Unknown	8	32	
Race			
White	274 (52)	368 (50)	
Black	236 (45)	323 (43)	
Other	19 (4)	49 (7)	0.06
Unknown	5	20	
CD4 cell count at delivery (cells/mm^3)			
Median (IQR)	418 (310-597)	413 (286-582)	0.33
<200	37 (9)	76 (12)	
200-499	230 (54)	345 (53)	
≥ 500	156 (37)	225 (35)	0.28
Unknown	111	114	
HIV RNA at delivery (log_{10} copies/ml)			
Median (IQR)	1.69 (1.69-5.56)	1.69 (1.69-2.60)	0.18
Unknown	136	145	
Parity (previous live births)			
None	239 (45)	338 (46)	
1	188 (36)	230 (31)	
2	71 (13)	108 (15)	
≥3	31 (6)	56 (8)	0.33
Unknown	5	28	
Injecting drug use during pregnancy			
No	511 (97)	728 (97)	
Yes	16 (3)	23 (3)	0.89
Unknown	7	9	

† n (%) unless otherwise stated

‡ *p* values were calculated with the Chi-squared or Mann-Whitney test as appropriate

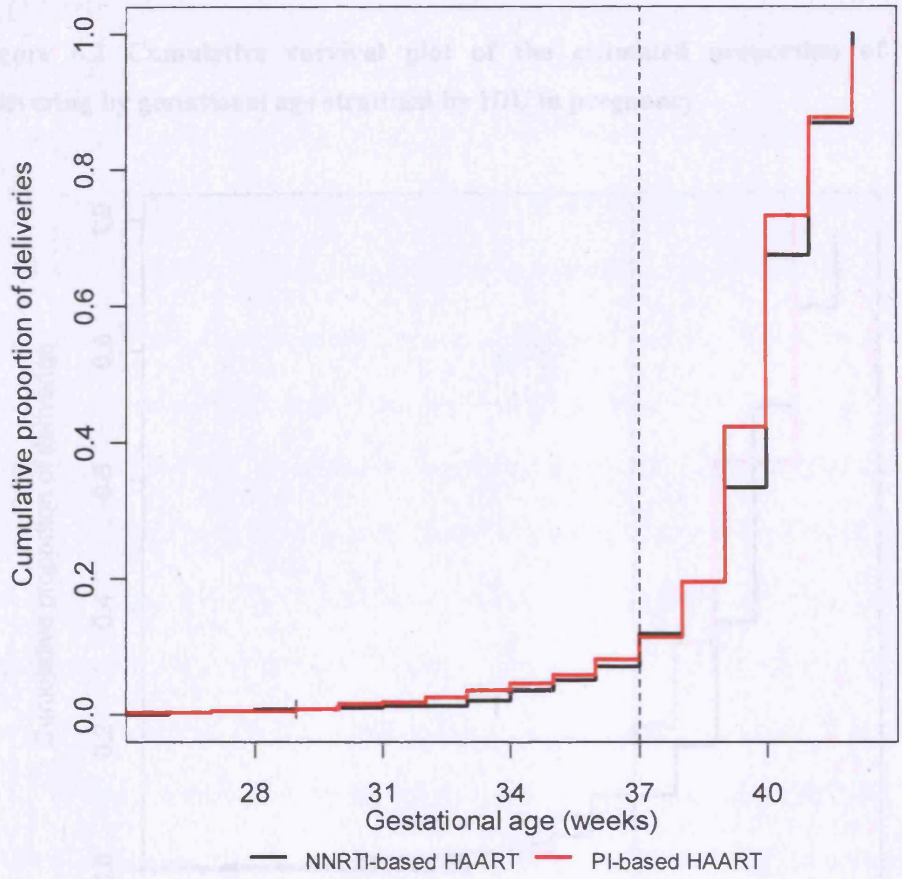
Table 6.2 contd.

Characteristic	NNRTI-based HAART (*n*= 534)	PI-based HAART (*n*= 760)	*p* value[†]
Initiation of ART			
Pre-pregnancy	156 (31)	217 (31)	
First trimester	57 (12)	68 (10)	
Second trimester	168 (34)	315 (45)	
Third trimester	115 (23)	98 (14)	<0.001
Unknown	38	62	
Year of delivery			
1995-1997	0	4 (0.5)	-
1998-2000	79 (15)	232 (30.5)	
2001-2003	313 (59)	335 (44)	
2004-2006	142 (27)	189 (25)	<0.001
Mode of delivery			
Vaginal	107 (20)	171 (23)	
Emergency CS	38 (7)	76 (10)	
Elective CS	388 (73)	512 (67)	0.08
Unknown	1	1	
Median gestational age at delivery (range)	38 (26-42)	38 (23-42)	0.11
≥ 37 weeks	405 (76)	605 (80)	
34-36 weeks	100 (19)	112 (15)	
<34 weeks	29 (5)	43 (6)	0.16

† n (%) unless otherwise stated

‡ *p* values were calculated with the Chi-squared or Mann-Whitney test as appropriate

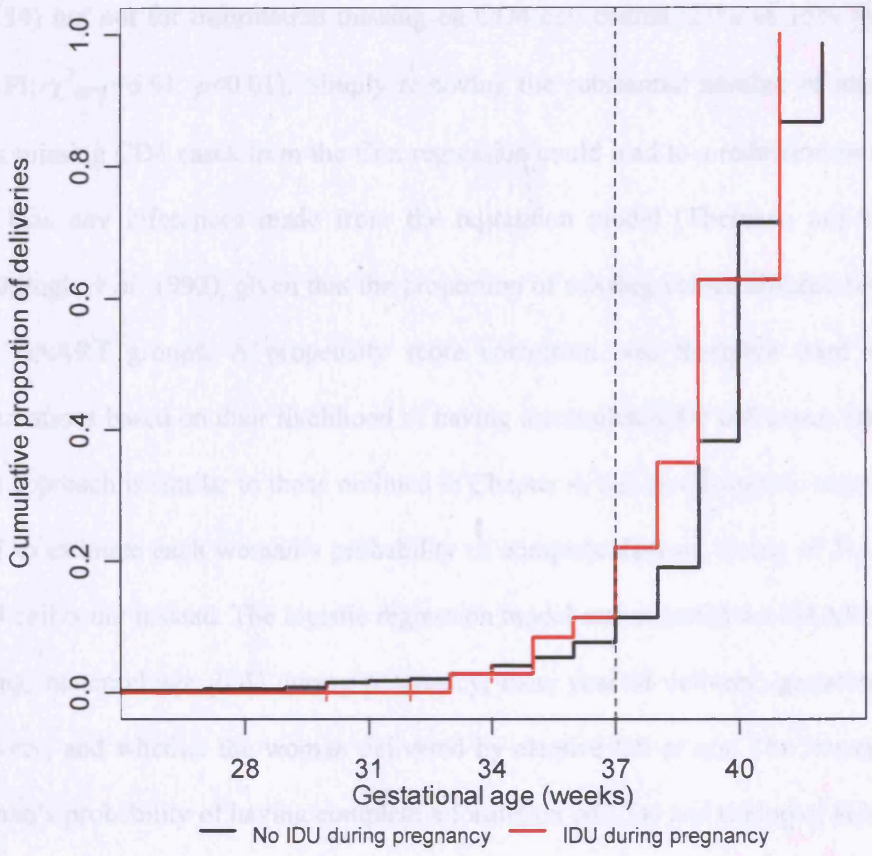
Figure 6.1 Cumulative survival plot of the estimated proportion of women delivering by gestational age stratified by HAART regimen



An estimated 3.6 % (95%CI 2.3-5.7) and 4.5% (3.3-6.3) of women receiving NNRTI- and PI-based HAART, respectively had delivered by 34 weeks, and 11.9% (95%CI 9.3-15.2) and 11.5% (95%CI 9.4-14.1), respectively, by 37 weeks. A log-rank test comparing the two HAART groups revealed no evidence for significant differences over gestational ages (χ^2=1.1, p=0.29). Log rank tests revealed no differences between the estimated time to delivery by maternal CD4 count ($\chi^2_{df=2}$=3.2; p=0.21), time of starting HAART ($\chi^2_{df=3}$=3.1; p=0.38), maternal age ($\chi^2_{df=2}$=2; p=0.38) and race ($\chi^2_{df=2}$=1.4; p=0.49). However the log rank test revealed significant differences between women who used injecting drugs in pregnancy and those who did not ($\chi^2_{df=1}$=4.9; p=0.03). A plot of the

associated survival curves shown in Figure 6.2 and reveal a higher proportion of deliveries for IDU.

Figure 6.2 Cumulative survival plot of the estimated proportion of women delivering by gestational age stratified by IDU in pregnancy



An estimated 4.0 % (95%CI 3.0-5.4) and 11% (95%CI 9.6-13.3) of women who did not use injecting drugs during pregnancy had delivered by 34 and 37 weeks respectively, compared to an estimated 2.7% (95%CI 0.4-18) and 22% (95%CI 11.0- 41.1)of women who did.

A Cox proportional-hazards model was used to investigate the association between antenatal HAART regimen and risk of premature delivery further while adjusting for

confounders. There were 364 women with information missing on the variables considered for inclusion in the model, with the majority of these cases missing information on maternal CD4 cell count or timing of treatment (Table 6.2). The proportion of women missing information on timing was similar between treatment groups (7% vs 8% for NNRTI-containing and PI-containing HAART; ($\chi^2_{df=1}$=0.34, p=0.34) but not for information missing on CD4 cell counts (21% vs 15% for NNRTI and PI; $\chi^2_{df=1}$=6.91, p<0.01). Simply removing the substantial number of mother-child pairs missing CD4 cases from the Cox regression could lead to a reduction in efficiency and bias any inferences made from the regression model (Therneau and Grambsch 2000;Pugh *et al.* 1992), given that the proportion of missing values differed between the two HAART groups. A propensity score correction was therefore used to weight observations based on their likelihood of having incomplete CD4 cell count information. This approach is similar to those outlined in Chapter 4, but here a logistic regression was used to estimate each woman's probability of complete data on timing of HAART and CD4 cell count instead. The logistic regression model was adjusted for HAART type and timing, maternal age, IDU during pregnancy, race, year of delivery, gestational age at delivery, and whether the woman delivered by elective CS or not. The inverse of each woman's probability of having complete information on CD4 and timing of HAART was then included as case weights in the Cox regression model. Conservative 95% confidence intervals for ARHs were derived using robust standard errors, an appropriate variance estimate for weighted Cox regression models (Therneau and Grambsch 2000).

The RHs and ARHs from the Cox regression model with no adjustment for missing information and WARHs from the weighted Cox regression model are given in Table 6.3.

There was no evidence of a violation of the proportional hazards assumption for the model as a whole (global test, p=0.57), or with any individual covariates in the model (all p>0.06).

In unadjusted analyses, the RH for delivery was similar between the two HAART groups. The only significant univariable predictors of the rate of delivery was the use of injecting drugs during pregnancy and having a parity of ≥3 live births (compared to no previous live births; Table 6.3).

Although the estimated ARHs and WARHs were roughly similar between the two Cox regression models, the standard errors were not and the associated p-values and 95% CIs in the weighted model were more conservative. The main difference was in the estimate for IDU in pregnancy. Both models estimated an approximate doubling in the rate of delivering for women using injecting drugs during pregnancy, but the effect was slightly higher and statistically significant in the model without the probability weights. By including information from the missing variables which were found to depend on treatment group, the weighted Cox regression model appeared to result in more conservative confidence intervals; as a result, the WARH estimate for IDU in pregnancy was slightly lower than the ARH but was not significant.

Table 6.3 Relative risks for delivery from weighted Cox regression model (n=930)

	Model without probability weights				*Weighted model*	
	RH (95% CI)	***p***	**ARH† (95% CI)**	***p***	**WARH† (95% CI)**	***p***
Antenatal HAART regimen						
NNRTI-based	1.00		1.00		1.00	
PI-based	1.10 (0.86-1.40)	0.46	1.10 (0.85-1.41)	0.47	1.12 (0.89-1.43)	0.33
Time of starting HAART						
Third trimester	1.00		1.00		1.00	
Second trimester	1.14 (0.83-1.57)	0.42	1.25 (0.90-1.75)	0.18	1.21 (0.88-1.67)	0.25
First trimester	1.03 (0.64-1.68)	0.90	1.16 (0.71-1.92)	0.55	1.11 (0.68-1.81)	0.68
Pre-pregnancy	1.12 (0.79-1.61)	0.52	1.22 (0.83-1.78)	0.31	1.20 (0.83-1.75)	0.34
Race						
White	1.00		1.00		1.00	
Black	0.97 (0.76-1.24)	0.81	0.94 (0.72-1.21)	0.62	0.96 (0.74-1.24)	0.73
Other	0.80 (0.47-1.37)	0.42	0.77 (0.44-1.34)	0.36	0.78 (0.46-1.30)	0.33
Maternal age						
≤25	1.00		1.00		1.00	
26-34	0.99 (0.75-1.33)	0.99	0.89 (0.66-1.22)	0.48	0.91 (0.68-1.23)	0.54
≥35	0.92 (0.66-1.29)	0.62	0.76 (0.53-1.10)	0.15	0.76 (0.52-1.09)	0.13
Use of IDU						
None/ex-user	1.00		1.00		1.00	
Used during pregnancy	2.06 (1.06-4.01)	0.03	2.17 (1.09-4.33)	0.03	1.99 (0.87-4.57)	0.10
CD4 cell counts at delivery						
>=500	1.00		1.00		1.00	
200-499	1.07 (0.84-1.37)	0.59	1.09 (0.84-1.41)	0.50	1.09 (0.85-1.40)	0.48
<200	1.18 (0.75-1.83)	0.48	1.23 (0.78-1.94)	0.38	1.19 (0.74-1.92)	0.48
Parity						
No previous livebirths	1.00		1.00		1.00	
1 livebirths	1.23 (0.94-1.61)	0.13	1.30 (0.98-1.72)	0.07	1.30 (0.99-1.73)	0.06
2 livebirths	1.29 (0.91-1.85)	0.16	1.45 (0.99-2.12)	0.06	1.46 (0.98-2.16)	0.06
≥3 livebirths	1.84 (1.23-2.74)	<0.01	2.18 (1.42-3.35)	<0.001	2.20 (1.50-3.23)	<0.001

† model adjusted for all other variables in the table

In the weighted Cox regression model, the estimated rate of women delivering was marginally higher for women receiving PI-based HAART than NNRTI-based HAART in pregnancy, but this was not statistically significant (WARH 1.12 [95%CI 0.89-1.43]). Additionally, there was no significant association between the rate of women delivering and the timing of HAART initiation, CD4 cell counts at delivery, maternal age, race or IDU (WARHs, Table 6.3). The WARH estimate for IDU in pregnancy suggested a doubling of the rate of delivering and although the 95%CI contained one, it suggested that IDU could be associated with as much as a four-fold increase in the rate however this study had only limited power to detect an effect owing to the small number of women IDUs. The rate of delivery for women with three or more previous live births however was estimated to be twice that of nulliparous women, and was the only significant predictor in the model.

Of clinical relevance are proportion of deliveries at 34 and 37 weeks of gestation; as proportional hazards are assumed in a Cox model (an assumption which was shown to hold for this model), the above interpretation of a similar rate of women delivering between HAART regimens, and an increased rate among women who have three or more previous live births also apply to the rate at 34 or 37 weeks of gestation.

6.3 Key points

- There was a similar pattern and prevalence of congenital abnormalities among infants exposed to antenatal ART and those who were not and this was also true for exposure to HAART.

- There was no evidence to suggest that exposure to first trimester ART increases the risk of congenital abnormalities.

- The proportion of women delivering by 37 weeks of gestation was around 11% for women receiving PI- or NNRTI-based HAART.

- The use of PI-based HAART was not associated with an increased rate of women delivering in adjusted analyses compared to NNRTI-based HAART.

- Women who have delivered three or more previous live births had a significantly higher rate of delivery than nulliparous women.

Chapter 7 Conclusions on the analysis of data on HIV infection in pregnancy

7.1 Introduction

HAART in resource-rich settings has substantially reduced MTCT rates through successful suppression of maternal HIV RNA (Mofenson *et al.* 1999;European Collaborative Study 2005b). An increasing proportion of HIV-infected pregnant women in these settings are identified and treated before pregnancy with a substantial minority diagnosed antenatally who start HAART for the first time in pregnancy (European Collaborative Study 2005b). In many Western European countries, these women are increasingly likely to have acquired HIV heterosexually and to be from countries with generalised epidemics, mostly sub-Saharan Africa. Data available on whether pregnancy may impact on HIV RNA and CD4 levels and the response of HAART, and whether any effect varies by the characteristics of these women, are limited.

This thesis explores the changes in CD4 counts and HIV RNA viral load during pregnancy in HIV-infected women treated with HAART from conception as well as changes in viral load in pregnancy in untreated women with the aim of informing understanding of the effect of pregnancy, race and IDU and the impact of ART on these markers. Viral response to initial HAART regimens in pregnancy was determined in treatment-naïve women and the association between antenatal ART and adverse pregnancy outcomes (birth defects and premature delivery) was assessed.

These findings using data collected within the ECS contribute to the evidence-base for management and understanding of HIV infection in pregnancy. Analyses were based on pregnancy data collected since 1987 from centres across Europe, which in recent years have been shown to be highly representative of the clinic population from which they were drawn.

7.2 Virological and immunological changes over pregnancy

The untreated and treated women in these analyses were all relatively healthy at the beginning of pregnancy with a median of 437 and 381 cells/mm^3, respectively. A healthy pregnant woman effect has been described elsewhere among HIV-infected women in resource-rich settings (Bessinger *et al.* 1998) and uninfected women in developing countries (Ronsmans *et al.* 2001). An American study of HIV infected women reported significantly higher CD4 cell counts among the 192 pregnant women in the study, compared to those that were not pregnant (median 519 vs. 333 cells/mm^3) (Bessinger *et al.* 1998). Possible explanations for this difference is that the pregnancy confers benefits to the women through the associated physiological changes or, more likely, that this simply represents a selection bias, where healthier women are more likely to become pregnant than those who are not (Ronsmans *et al.* 2001).

The few studies that have addressed HIV RNA dynamics in pregnancy in untreated HIV-infected women, have reported levels to be relatively stable over pregnancy (Burns *et al.* 1998;Mayaux *et al.* 1997a;O'Shea *et al.* 1998), or described only small and non-significant decreases in viral load (Cao *et al.* 1997;Watts *et al.* 2003;Melvin *et al.* 1997). However, only two of these five studies used appropriate statistical methods for correlated longitudinal data (Burns *et al.* 1998;Watts *et al.* 2003), and most did not

report adjustment for maternal immune status (Cao *et al.* 1997;O'Shea *et al.* 1998;Watts *et al.* 2003;Mayaux *et al.* 1997a;Melvin *et al.* 1997), which together with HIV RNA levels is an important prognostic indicator (Mellors *et al.* 1997). Here, HIV RNA viral load during pregnancy was assessed for untreated women in the ECS using a linear mixed effects model, which was able to account for variability in viral load measurements within and between women and also allowed for women with only one measurement to contribute to estimation of the model (Chapter 3). This analysis included 324 women with viral load measurements available from the start of the second trimester of pregnancy, a relatively larger number of exclusively untreated women than included in previous studies. After adjusting for CD4 count status and race, HIV RNA levels did not remain stable during pregnancy: white women were estimated to have a 4.3% weekly decrease in viral load from the start of the second trimester up to the time of delivery. For black women, mostly born in sub-Saharan Africa, the change in RNA viral load was estimated to be positive and significantly different to the change in viral load for white women. However, the increase in viral load over pregnancy in black women was not significantly different to zero, i.e. the slope for viral load in pregnancy was relatively stable for black women and decreasing for white women.

These results stand in contrast to published studies of untreated women in pregnancy that report no significant association between gestation and viral load in untreated HIV-infected pregnant women, or those receiving mono or dual therapy (Cao *et al.* 1997;O'Shea *et al.* 1998;Watts *et al.* 2003;Mayaux *et al.* 1997a;Melvin *et al.* 1997). A study of 204 women in the Ariel project reported that viral load did not vary significantly during pregnancy (estimated decrease of 2 copies/ml/day) for 29 untreated women or those receiving ZDV monotherapy in unadjusted linear regression, but rose

significantly in the postpartum period (Cao *et al.* 1997). An analysis using pairwise comparisons (with a two-sample Wilcoxon test) of viral load values involving 254 pregnant women in the SEROGEST cohort without adjustment for CD4 count, revealed no significant variations during pregnancy among, although postpartum values were significantly lower (Mayaux *et al.* 1997a). Women enrolled in this French study had a median viral load in the second trimester of pregnancy of 3.5 $\log_{10}$ copies/ml and an overall mean CD4 count of 529 cells/mm^3 (Mayaux *et al.* 1997a), similar to the median viral load and CD4 count in untreated women in the ECS of 3.51 $\log_{10}$ copies/ml and 437 cells/mm^3 (Chapter 3). None of these studies used an appropriate statistical method for correlated longitudinal data, which could explain the discordance in results seen between these studies and in untreated women in the ECS. Using data on 160 women enrolled in the US Mother and Infants Cohort Study mostly from the third trimester of pregnancy, an LME model adjusting for blood sample type did not reveal any decreases in viral load in pregnancy, and revealed only increases in the second year postpartum; IDU and use of ZDV during pregnancy was not associated with a significant change in viral load and was therefore not included in the final model (Burns *et al.* 1998). In a sub-analysis of the 48 women with at least two HIV RNA measurements from before the third trimester and during the third trimester (mostly close to delivery), there was no overall change in viral load levels in pregnancy (Burns *et al.* 1998). However, although this US study was able to account for correlated measurements, it included only 48 women with limited measurements available from before the third trimester of pregnancy and the limited duration of follow-up in pregnancy may explain why relatively stable levels of viral load were seen in pregnancy in this study but not in the analysis carried out in this thesis (Burns *et al.* 1998). In contrast to the first two studies mentioned above, the analysis in Chapter 3 used an LME model to account for repeated

measures and unlike the analysis in the Mother and Infants Cohort Study, was also able to adjust for immune status in pregnancy, which was strongly associated with levels of viral load; women with CD4 counts <200 cells/mm^3 were estimated here to have a 6-fold higher viral load than women with CD4 counts ≥500 cells/mm^3, highlighting the importance for adjustment for this variable.

There were 367 measurements on 246 untreated women available in the ECS analysis of viral load in pregnancy, larger than the numbers included in the studies above with the exception of the SEROGEST cohort, which included data on 526 measurements on 254 untreated women in pregnancy (Mayaux *et al.* 1997a); the analyses in both studies included women with one measurement only. The majority of untreated women in the ECS with only one measurement, had this taken at the time of delivery (Chapter 3) and although this contributed to estimation of the intercept, it will have added only limited information to the estimate of the slope, which may be confounded by factors influencing when in pregnancy this measurement was made. However, a sensitivity analysis including only the 80 women with at least two measurements (n=215) also revealed a decrease in HIV RNA viral load over pregnancy albeit of smaller magnitude, which suggests that the influence of women with single measurements on the estimate of the slope in pregnancy was not large. The influence of these measurements on the intercept estimate however was greater, which was estimated to be ~0.4 $\log_{10}$ copies/ml larger when excluded, reflecting the fact that they were taken around the time of delivery/intercept and with generally lower levels of HIV RNA (Chapter 3). In addition, the analysis in the SEROGEST study included only paired comparisons of values between two time points, whereas through the use of an LME model the analysis in this thesis was able to include all measurements across the second and third trimester of

pregnancy while adjusting for factors associated with the level of viral load; the differences in statistical methods may help explain why a significant decrease in pregnancy was seen in the untreated women in the ECS, but not in this French study.

No studies to date have examined levels and patterns of viral load during pregnancy among women receiving HAART, with most focusing on viral load at the time of delivery and the associated risk of MTCT (Cooper *et al.* 2002;Dorenbaum *et al.* 2002;Mofenson *et al.* 1999). In the analyses presented in this thesis, among the group of women on HAART at conception and with detectable viral load at this time, estimated changes in HIV RNA viral load in the first trimester of pregnancy were dependent on CD4 count and thereafter decreased by 14% per week up to the time of delivery, independent of race and history of IDU, CD4 count and type of HAART regimen (Chapter 5). This group of treated women had similar viral loads at the beginning of pregnancy to untreated women in the ECS (median in first trimester of 3.50 $\log_{10}$ copies/ml and 3.52 $\log_{10}$ copies/ml for untreated and treated women, respectively). The decrease in HIV RNA viral load from the second trimester of pregnancy to delivery was almost four times as rapid among these treated women as observed in untreated women over the same period in pregnancy, reaffirming the effectiveness of HAART in reducing viral load in pregnancy. Recent data on viral load changes in previously untreated adults have shown a very rapid decrease in the first six weeks of treatment, with HIV RNA levels decreasing by 97% per month during this period, with a considerably more gradual decline, of 14% each year thereafter (Thiébaut *et al.* 2006). This group of treated women in the ECS had all been on HAART for longer than 1.5 months by the start of the second trimester and although a very different group, the magnitude of the

decreases seen here in pregnancy may be greater than would be expected from treatment alone.

The finding of a decrease in HIV RNA viral levels over pregnancy in untreated white women (Chapter 3) and a relatively large decrease in treated women (Chapter 5) suggest a possible mechanism related to physiologic changes during pregnancy. Haemodilution occurring in pregnancy results in increases in plasma volume of up to 50%, with most of the increase in volume occurring after the first trimester (Hytten 1985). The likelihood-based time of change-point for the viral load and CD4 models was determined to be around the beginning of the second trimester and the two slope change-point models shown in Chapter 5 may therefore reflect this physiological change (Hytten 1985). In the study of untreated ECS women, viral load was estimated to decrease by 0.019 $\log_{10}$ copies/ml per week, roughly equal to a decrease of 0.46 $\log_{10}$ copies/ml in the 24 weeks of follow up over the second and third trimesters of pregnancy; this equates to a decrease of 65% and therefore may not be completely explained by changes in plasma volume alone. Plasma HIV RNA levels have been reported to be relatively stable on a week-to-week or month-to-month basis in clinically stable non-pregnant patients, as long as antiretroviral therapy is not initiated or changed (Saag *et al.* 1996). Intra-assay sample variability of HIV RNA is between 0.1 and 0.2 $\log_{10}$ copies/ml and the natural biologic variation of HIV RNA in plasma is around 0.3 $\log_{10}$ copies/ml (Saag *et al.* 1996). In practice, therefore a change in viral load of > 0.5 $\log_{10}$ is generally considered to reflect a biologically relevant change in the level of viral replication. Although the mean decrease of 0.46 $\log_{10}$ copies/ml is small enough to be consistent with biological and intra-assay variability, the confidence interval for the change over pregnancy was between -0.17 and -0.72 $\log_{10}$ copies/ml indicating that the data may also be consistent

with a decrease large enough to indicate a possible pregnancy effect. Immunological changes occurring in pregnancy including the depression of cell-mediated immunity and a shift to a type 2 (humoral immunity) cytokine environment, have been shown to be important in modifying HIV replication and may impact on levels of CD4 counts (see below also), as well as those of HIV RNA viral load (Weinberg 1984;Kidd 2003;Johnstone *et al.* 1994). More specifically, pregnancy is associated with increased levels of particular cytokines, such as interleukin-10, which have been reported to down regulate HIV-replication and may explain the decreasing levels of HIV RNA in pregnancy seen here, over and above that associated with biologic and assay variability alone (Montaner *et al.* 1994).

In a sub-analysis of untreated women who had at least two HIV RNA measurements in pregnancy, the correlation of $\log_{10}$ HIV RNA within women was obtained from the LME model and was estimated to be 0.66. This estimate was similar to the correlation estimate of 0.69 for 198 men in the MACS, who contributed 1139 HIV RNA measurements during the period of minimal antiretroviral use (Yamashita *et al.* 2003). Although based on comparatively small numbers per individual, this result suggests that gestation over the second and third trimester of pregnancy does not have a large effect on the within-woman correlation of viral RNA measurements, when compared to that observed in HIV-infected men.

The analysis of CD4 counts in treated pregnant women in the ECS revealed levels to decrease in early pregnancy and to increase thereafter (Chapter 5). However, after adjusting for viral load category, CD4 counts were estimated to be stable in early pregnancy and to increase by 0.32% per week thereafter, although this increase did not

reach statistical significance. In both HIV-infected and uninfected women, findings regarding pregnancy-related changes in immunological markers conflict, with reports of declining, stable and increasing CD4 cell counts over pregnancy (Johnstone *et al.* 1994;Brettle *et al.* 1995;Castilla *et al.* 1989;van Benthem *et al.* 2005). A study from the Mother and Infants Cohort Study comparing changes in CD4 levels during pregnancy and postpartum in 340 HIV untreated seropositive and seronegative women estimated CD4 counts to decline by 0.68% per month during pregnancy (Burns *et al.* 1996). The non-significant increase in CD4 counts observed in this analysis of treated women from the ECS therefore may represent a tempering of treatment-related increases in CD4 counts due to a naturally occurring decline in pregnancy. A recent study of 198 HIV-infected pregnant women in Ireland found that women receiving treatment prior to conception had a significantly reduced decline in CD4 counts between pre-conception and the first trimester (mean 485 to 410 cells/mm^3), compared to women who did not receive treatment during pregnancy (mean 575 to 421 cells/mm^3) and may be due to a similar mechanism in pregnancy (Mucahy *et al.* 2006). A pregnancy-related decline has also been supported by findings of an increase in CD4 counts postpartum in untreated and treated women (Mucahy *et al.* 2006;Ekouevi *et al.* 2006;ECS and Swiss Cohort and European Collaborative Study and the Swiss HIV Pregnancy Cohort 1997). The influence of pregnancy on immunologic changes is poorly understood, and these pregnancy-related changes in CD4 could be due to a variety of factors, such as haemodilution in pregnancy (Johnstone *et al.* 1994;Brettle *et al.* 1995;Ekouevi *et al.* 2006). It has also been suggested that the decline in CD4 in early pregnancy may be partially responsible for the survival of the fetal allograft (Johnstone *et al.* 1994). An alternative explanation is that the immunological changes observed could be due to recompartmentalisation of CD4 cells in response to hormonal changes in pregnancy

(Johnstone *et al.* 1994). This is the first study to examine changes in CD4 counts over pregnancy in HAART-treated women and revealed some increases in CD4 count over the last two-thirds of pregnancy. Adjusting for viral load category at the time of the CD4 measurement improved the reliability of the estimates and only the highest viral load category ($\geq$3.25 $\log_{10}$ copies/ml) was a significant predictor of mean CD4 count, with CD4 counts estimated to be 15% lower in this group than those in women with viral loads <2.30 $\log_{10}$ copies/ml. The modified estimates of the two markers after adjustment for CD4 and viral load (Chapter 5) confirm the well established relationship between the two markers (Mellors *et al.* 1997;Boscardin *et al.* 1998;Thiébaut *et al.* 2002) in pregnancy. Unlike other studies which found a significant correlation between the slopes and intercepts of viral load and CD4 from bivariate models in patients initiating HAART (Thiébaut *et al.* 2005;Thiébaut *et al.* 2006;Thiébaut *et al.* 2003), there was only evidence of a negative correlation between the intercepts for the two markers.

The time period of delivery variable (1998-1999, 2000-2001 and 2002-2006) was included in the analysis of viral load among women receiving HAART at the time of conception to account for changes in HIV management, although its parameter estimates were not statistically significant (Chapter 5). However, this may have been due to limited statistical power and the parameter estimates and 95% CIs suggest that a trend of lower HIV RNA levels over the time periods cannot be ruled out. This trend could be representative of general improvements in HIV care including better adherence and health of the pregnant women and was evidenced also by an increase in the number of undetectable measurements over these time periods. However, although a conditional-likelihood method was used to take account of left-censoring, HIV RNA assay specific cut-offs were not adjusted for as this data was not routinely collected in

the ECS, and improvements in the assay detection limits over the time periods may also partially explain this finding.

7.3 Choice of initial HAART regimen in pregnancy and time to achieving an undetectable HIV RNA viral load by delivery

Suppressing plasma HIV RNA viral load below detectable limits is one of the goals for effective management of HIV-infected women in pregnancy and for PMTCT (British HIV Association 2005;Public Health Service Taskforce 2006a). Among the analysis of previously ARV-naïve women in the ECS initiating HAART during pregnancy, most (73%) delivered with undetectable viral load and the remainder delivered with detectable, but generally very low levels (Chapter 4). Less than a quarter of women had immunological indications for treatment (British HIV Association 2005;Public Health Service Taskforce 2006a), with the remaining women started on HAART primarily for PMTCT.

Most women were prescribed PI-containing HAART, with a highly homogenous approach, with 76% of these women on NFV+ZDV+3TC. A third of women received NVP-containing HAART, with increasing use over time (i.e. between the periods 1997-2000 and 2003-2004). Updated NVP prescribing advice, which recommends that NVP be avoided in women with CD4 cell count above 250 cells/mm^3 because of the potential increased risk of hepatotoxicity (Public Health Service Taskforce 2006b), suggests that this trend is unlikely to continue: only 12 women in the ECS initiated NVP with CD4 counts above the recommended threshold in 2005 (unpublished data). The predominance of ZDV+3TC-containing regimens here reflects current recommendations for this NRTI combination as the backbone for pregnant women

(DHHS Panel on Antiretroviral Guidelines for Adults and Adolescents 2006;Public Health Service Taskforce 2006a) and is consistent with prescribing patterns in Europe in non-pregnant individuals (Mocroft *et al.* 2006b).

Adjusting for baseline prognostic factors, the hazard of achieving undetectable viral load in this cohort was greater for women initiating HAART with NVP, with women initiating with PI-based HAART taking on average 1.4 times longer to achieve viral suppression. To my knowledge, this analysis is the first to suggest that choice of initial HAART regimen has implications for timely achievement of undetectable viral load in pregnancy. Accumulating evidence from adult cohorts suggest that types of initial HAART may have a differential impact on the time to short term immunological and virological response and delaying disease progression, particularly the long-term sustainability of viral suppression and immune reconstitution (Matthews *et al.* 2002;Mocroft *et al.* 2006b;Mocroft *et al.* 2006a;The Antiretroviral Therapy Cohort Collaboration 2006). A study of 789 patients from EuroSIDA found that those receiving HAART with a single PI (mostly NFV) were significantly less likely to achieve undetectable viral load <500 copies/ml (RH 0.74) than women receiving NNRTI-based HAART (mostly efavirenz) (Mocroft *et al.* 2006a). A London study of 888 patients also reported a differential response between regimens, with the rate of achieving undetectable viral load significantly higher for patients on efavirenz than those on a single PI or NVP (Matthews *et al.* 2002). However, a Canadian population-based analysis of 439 ART-naïve HIV-infected patients, most of whom were women, found that the use of NNRTI-containing HAART (94% NVP) was associated with a more rapid HIV-RNA suppression compared to both patients receiving double PI or single PI-

containing HAART in adjusted analyses, which is in line with the findings presented in this thesis (Wood *et al.* 2003).

Findings from randomised controlled trials in non-pregnant adults conflict on the relative effectiveness of PI-(non-boosted) versus NNRTI-containing HAART. A recent direct meta-analysis of "head-to-head" randomised trials suggested NNRTI-based HAART (predominated by efavirenz) was 60% more effective for virological suppression than PI-based HAART (50% boosted PIs), although no difference in clinical outcomes was reported (Chou *et al.* 2006). However, an indirect meta-analysis yielded contradictory results, i.e. NNRTI-based HAART was worse than PI-based HAART for virological suppression (Chou *et al.* 2006); these discordant results may be due to differences in population, study design or type of NRTI backbone, highlighting the difficulties in translating trial findings into clinical recommendations and the importance of direct comparisons (Chou *et al.* 2006). The comparison of NVP-based HAART and PI-based HAART in the analysis here was facilitated by the overall similar NRTI backbone used in conjunction with these components.

Therapeutic decision-making in pregnancy is complicated by unique factors including the need to consider PMTCT, safety/toxicity and physiological changes, which may affect pharmacokinetics (Public Health Service Taskforce 2006b;Mirochnick and Capparelli 2004). Accumulating data on NFV pharmacokinetics suggest that drug levels in the third trimester may frequently be sub-therapeutic (Villani *et al.* 2006;Nellen *et al.* 2004;van Heeswijk *et al.* 2004). This may explain the finding of a superior virological response to women receiving NVP-based HAART, in contrast to the Combine Study, where an equivalent response was reported among ARV-naïve non-pregnant individuals

randomised to a ZDV+3TC backbone with NVP or NFV (Podzamczer *et al.* 2002). A sensitivity analysis examining the time taken to achieve undetectable viral load between NFV-based and NVP-based HAART also revealed similar results of a superior virologic response with NVP-containing HAART, reinforcing the possibility of underlying differences in pharmacokinetics explaining the differing responses among the two treatment groups.

There were no significant differences in time to attaining undetectable viral load between severely immunosuppressed women (CD4 count <200 cells/mm^3) and those with greater immunocompetence. This is consistent with other studies which have found a significant association between baseline viral load and subsequent virological response after HAART initiation, but not with baseline CD4 counts (Phillips *et al.* 2001;Matthews *et al.* 2002). Few studies have examined this association in ARV-naïve pregnant women and these findings suggest that the relationship between baseline CD4 count and viral load and subsequent treatment effect in this group of women may not differ from non-pregnant adults.

7.4 Adverse pregnancy outcomes

7.4.1 Prematurity

The association between HAART and prematurity has been consistently reported in Europe since 1998 (ECS and Swiss Cohort and European Collaborative Study and the Swiss HIV Pregnancy Cohort 2000;European Collaborative Study 2004;Townsend *et al.* 2007b). The reference groups in these studies assessing the risk of prematurity were mostly either untreated women, or those receiving mono or dual therapy. However, the relevance of this comparison group is questionable given that most women in resource-

rich settings are now expected to receive HAART, with either an NNRTI or PI component. These combinations have predominated in the ECS in recent years and other pregnancy cohorts have reported similar uses of these combinations in pregnancy (Townsend *et al.* 2007b).

In a comparison of PI-based or NNRTI-based HAART with an appropriate statistical analysis to account for the high number of elective CS deliveries, there were no differences between the two HAART groups with respect to the rate of women delivering by 34 or by 37 weeks gestation (Chapter 6). Adjustment for maternal IDU, race and maternal CD4 cell count, established risk factors for prematurity (Patel *et al.* 2004;Slattery and Morrison 2002;The European Collaborative Study 1996) increased the reliability of the estimated relative hazards. The only significant predictors of the rate of delivering in the final weighted and adjusted Cox model was the category of women with at least 3 previous live births, however IDU was significantly associated with a doubling of the rate of delivering in the adjusted model which did not include the probability weights. Although IDU was not statistically significant in the weighted model, the RH for IDU was similar and had a 95% CI which contained zero but was not clustered closely around zero; given the width of the 95% CI and the confirmed association between IDU in pregnancy and prematurity (Slattery and Morrison 2002), a finding of a non-significant result here suggests that this study was limited by statistical power to examine the effect of IDU.

A recent study of 4445 mother-child pairs exposed to ART and enrolled in the UK National Study of HIV in Pregnancy and Childhood (NSHPC) also found an equivalent risk of prematurity between those receiving HAART with PI compared to HAART

without a PI (including some women who received triple NRTI HAART) (Townsend *et al.* 2007b). A recent study of 681 women from Latin America and the Caribbean reported similar results (Szyld *et al.* 2006). However, these findings were in contrast to Cotter et al, who reported a 1.8 increased odds of premature delivery with PI-based HAART (single and boosted) compared to both monotherapy and HAART without a PI (exact details of the PI and NNRTI components were not given), in a US study of 999 women (Cotter *et al.* 2006). There are several possible explanations for the discrepancy in results between this ECS analysis and those from the Cotter study. Despite adjustment for several known confounders, data on other risk factors for prematurity such as prior preterm births, smoking and alcohol use in pregnancy, clinical disease stage and presence of sexually transmitted diseases were not routinely collected in the ECS and therefore could not be adjusted for (Slattery and Morrison 2002). Cotter et al. were able to adjust for these factors which could improve the reliability of their estimates, but most of these factors were reported not to have been significant factors of preterm delivery (odds ratios for these factors were not reported) (Cotter *et al.* 2006). However, the association between a history of a prior preterm was high and an adjusted odds ratio of 3.4 (95% CI 2.3-5.0) was reported. The inability to control for prior preterm deliveries in this ECS analysis may have resulted in a potential bias, especially if there was an imbalance in the number of women with a history of preterm births between the two treatment groups.

In the ECS analysis, there were no significant differences in the distributions of CD4 cell counts between treatment groups (Chapter 6). Cotter et al. described prescribing patterns in their cohort such that PI-based regimens were reserved for women who were immunocompromised and/or who had more advanced HIV disease (Cotter *et al.* 2006).

It could be possible therefore that the finding of an association between PI-based HAART and prematurity in this US study was confounded by indication for therapy of maternal disease. Furthermore, another principal difference between the studies is that the study by Cotter et al. used data from only one clinical site where the ECS analysis was based on data from up to 23 centres, where centre-associated differences were not accounted for directly. Standardised management protocols were used in both studies but the possibility of unmeasured centre-based differences is expected to be greater in the ECS and offers a further explanation for the discordance in these results. Although it is not possible to make any firm conclusions based on the ECS data given here, based on these data there was little evidence to suggest that the use of PI-based HAART regimens were associated with an increased risk of prematurity compared to the use of NNRTI-based HAART.

Finally, the overall crude severe prematurity rate was 5.6%, but there were no differences observed between women receiving PI-based or NNRTI-based HAART. Severely premature infants are likely to pose a substantial demand on health care services, both in the immediate and longer-term (European Collaborative Study 2004;Slattery and Morrison 2002).

7.4.2 Congenital abnormalities

In the ECS, 1.5% of children exposed to ART in utero were born with a congenital abnormality (Chapter 6). There was a similar pattern and prevalence of congenital abnormalities among infants exposed to antenatal ART and those who were not and this was also true for exposure to HAART. Furthermore, based on these data there was no evidence to suggest that exposure to first trimester ART increases the risk of congenital

abnormalities. As such, these findings are consistent with those of the Antiretroviral Pregnancy Registry (Covington *et al.* 2004;Antiretroviral Pregnancy Registry Steering Committee 2004). However, as a long-running birth-cohort study, the ECS benefits from a large number of mother-child pairs (both exposed and non-exposed), and are not subject to the potential for ascertainment and reporting bias which limit interpretation of registry data (Antiretroviral Pregnancy Registry Steering Committee 2004), and these findings are consistent with other observational studies from Europe (Townsend *et al.* 2006) (I. Grosch-Wörner, Personal communication).

7.5 Race and injecting drug use

Significant and substantial differences between black and white women were observed in the analysis of HIV RNA viral load and CD4 counts over pregnancy (Chapter 3 and 5) and in that investigating time to undetectable viral load in ARV-naïve pregnant women (Chapter 4). In black untreated women who were mostly born in Africa, the estimated change in viral load over pregnancy was significantly different to that of white women (mostly born in Western Europe), but this slope was not significantly different to zero. The estimated mean viral load at delivery for black women was 3.86 $\log_{10}$ copies/ml, and was significantly higher than white women who had a mean of 3.41 $\log_{10}$ copies/ml. This significant difference between HIV RNA levels conflicts with other European studies of pregnant and non-pregnant adults which have found black ethnicity to be associated with lower viral loads (Mayaux *et al.* 1997a;Saul *et al.* 2001). A longer duration of infection in these African women is a possible explanation for this difference, which has been associated with higher HIV RNA levels (Sabin *et al.* 2000;Lyles *et al.* 1999); another is the presence of co-infections. A recent ECS sub-study analysis of 1,050 women found that women born in Africa were at a significantly

increased risk of having a viral sexually transmitted infection (STI) in pregnancy (Landes *et al.* 2006). Concurrent co-infection with viral STIs, particularly the herpes virus, have been shown to increase the rate of viral replication, purportedly through increased levels of certain cytokines or through modulation of immune responses that control HIV viraemia (Anzala *et al.* 2000). The effect of haemodilution seen among white women therefore may have been attenuated by an increase in HIV RNA levels over pregnancy among black women and may explain the relatively stable slope of viral load in pregnancy, with significant differences seen at delivery only.

In the analysis of treated pregnant women (Chapter 5), white women with a history of IDU were estimated to have an almost 60% significantly lower HIV RNA viral load than white women with no history, but a similar slope of viral load decrease over pregnancy. A recent analysis by the CASCADE Collaboration based on 943 patients, most of whom were male (82%), found that the short term virological response was poorer in men infected by IDU than for men who have sex with men, but reported no differences in the long-term response (Thiébaut *et al.* 2006). Other studies have reported a negative effect of current IDU on virological and immunological responses to HAART (Lucas *et al.* 2001;Palepu *et al.* 2003). It has been suggested that the poorer response seen in IDUs is due to incomplete adherence (Spire *et al.* 2002), which does not appear be the case in this group of pregnant women who had a similar response to treatment regardless of IDU status but with lower initial levels of virus in those with IDU history. Of note, most of the women with a history of IDU were ex- rather than current users and thus would not be experiencing the chaotic lifestyle of an active IDU that is held to be the reason for poor adherence.

Among women treated with HAART from conception, there was no evidence of significant differences between white and black women with respect to viral load, although the estimated mean level was lower in black women (Chapter 5). Black women experienced similar changes in CD4 counts over pregnancy to white women, but were estimated to have 25% lower CD4 counts throughout pregnancy. This is consistent with a previous ECS analysis, which reported untreated black women to have 45% lower CD4 count compared to white women (Thorne *et al.* 1995). Possible explanations for the lower CD4 counts in black women, most of who were born in Africa, may include a longer duration of infection, or a higher prevalence of bacterial, protozoan and helminthic infections, which have been associated with lower CD4 cell counts (Lepage *et al.* 1993;Anzala *et al.* 2000;Chirenda 1999). Lower initial levels of CD4 cell count (i.e., pre-seroconversion) are another explanation, which may also be caused by a higher prevalence of bacterial infections in early life (Bunders *et al.* 2006). A study of HIV-uninfected children with or without in utero HIV exposure in Europe and Uganda found CD4 counts to be lower in Ugandan-born children than in children living in Europe with Ugandan mothers (Bunders *et al.* 2006). It has been suggested that differences in diet, environment and exposure to micro-organisms may be responsible for these differences and may also explain the differences seen in viral load and CD4 counts between black and white women in the ECS (Bunders *et al.* 2006;Ozkan *et al.* 1993).

In the analysis of time to undetectable viral load, black women responded to HAART more favourably than non-black women in unadjusted analysis; further investigation, stratifying by region of birth, revealed that this effect was limited to women of West African origin. African and non-African pregnant women had similar baseline immune

and virological status in this analysis; this is in contrast to previous findings based on the whole cohort (European Collaborative Study 2006) and those presented here in treated women (Chapter 5), in which black African women had lower CD4 counts at baseline. This probably reflects eligibility criteria for this analysis, i.e. all women were ARV-naïve and facilitated the comparison of treatment effect between these women. The median baseline CD4 count among black women here was marginally lower than those reported from African PMTCT trials, which were in the range of 335-363 cells/mm^3 in ARV-naïve pregnant women (Taha *et al.* 2006;Shapiro *et al.* 2006). Limited information is available regarding response to HAART in African populations, and even less in pregnant African women. In the DREAM pilot study in Mozambique, 65% of 40 pregnant women starting on HAART with median baseline HIV RNA 4.2 $\log_{10}$ copies/ml achieved viral suppression to <400 copies/ml by delivery, after an average of 12 weeks (Giuliano *et al.* 2006), consistent with the results of this ECS analysis (Chapter 4). An impact of race on disease progression or response to HAART has been suggested by several studies in pregnant and non-pregnant individuals, generally showing poorer virological responses in black and/or African groups, which have been suggested to be due to coinfections and/or adherence (Frater *et al.* 2002;Cunningham *et al.* 2004;Anastos *et al.* 2000). Data on adherence was not available in the ECS, but it seems unlikely that differing adherence levels could explain these findings, as the better virological response to HAART was limited to the West African group only. ARV drugs are usually developed using subtype B, the predominant subtype in resource-rich settings, as a reference virus, and little is known about the susceptibility of the different non-B HIV-1 subtypes to ARV drugs (Bocket *et al.* 2005;Geretti 2006). A study of 416 ARV-naïve patients infected with B and non-B HIV-1 subtypes found no differences between these groups with regards to the time to

achieving undetectable viral load in adjusted analyses (Bocket *et al.* 2005). A recent paper reviewing studies examining the association between HIV-1 subtype and the risk of disease progression and treatment response found that most studies were characterised by small numbers and short follow-up, and did not account for potential confounding such as the clustering of subtypes within risk and ethnic groups (Geretti 2006). The authors concluded that based on the limited clinical evidence, responses to HAART did not appear to differ significantly among subtypes, but further research will be necessary to confirm this (Geretti 2006). Immunogenetic and other human genetic variations have been known to impact on the natural history of HIV disease progression where individuals are untreated, but it is unclear whether these differences are relevant in the context of individuals receiving HAART (Brumme and Harrigan 2006). Differences in underlying maternal biological and genetic differences between Western African and European women may however explain the findings in this cohort (Chapter 4) (Brumme and Harrigan 2006). Another explanation is that West African women had better rates of adherence than Non-African women which would result in a faster response to treatment (Conway 2007;Moore *et al.* 2005), however adherence data were not collected in the ECS and therefore the results may be biased. Further work on treatment responses in African women needs to be carried out with adherence data collected to clarify these findings.

7.6 Management of HIV-infected pregnant women – implications of the research findings for clinical practice

In Western Europe, the combination of high antenatal coverage, universal HIV testing in most countries and access to a range of PMTCT interventions has resulted in very low MTCT rates of <1-2% (Townsend *et al.* 2007a;European Collaborative Study

2006). Maternal HIV RNA viral load is the most important predictor and modifiable risk factor for MTCT (European Collaborative Study 1999;Cooper *et al.* 2002;Ioannidis *et al.* 2001;Mofenson *et al.* 1999). Viral load is also a strong predictor of progression to AIDS and death (Mellors *et al.* 1997;Sabin *et al.* 2000). Therapeutic decision-making and management in pregnancy must consider the risk of transmission to infants by maintaining viral suppression to the time of delivery, but also the longer-term sustainability of treatment started or maintained in pregnancy. These decisions are further complicated by unique factors in pregnant women including the need to consider maternal and fetal safety, potential adverse pregnancy outcomes and physiological changes during pregnancy, which may affect pharmacokinetics (Public Health Service Taskforce 2006b;Mirochnick and Capparelli 2004).

With the widespread use of HAART in developed settings and increasing number of women having subsequent pregnancies, a large number of infected pregnant women can be expected to be receiving HAART before pregnancy (European Collaborative Study 2005b;Boer *et al.* 2007). Sequential measurements of viral load and CD4 cell counts are crucial for monitoring treatment response in HIV-infected individuals and help inform clinical decision-making, particularly regarding switching regimens. The findings here facilitate the interpretation of changes in these markers among HIV-infected pregnant women receiving HAART at conception, and may be useful for examining treatment response in pregnancy. However, it should be noted that these women were a select group in that they all had detectable viral loads at their first visit in pregnancy. These results may therefore not be generalisable to women who begin pregnancy on HAART with undetectable viral loads and should be interpreted with this in mind.

In resource-rich settings where HAART is commonplace, the added benefit of elective CS for PMTCT has been questioned for women with undetectable HIV RNA viral loads (Boer *et al.* 2007;British HIV Association 2005). In particular, the risk of vertical transmission without elective CS has been reported to be as low as 0-2% and there are concerns that the benefits of this intervention may be outweighed by the costs, such as increased risk of post-partum complications and greater burden on health care services associated with elective CS (Boer *et al.* 2007;Marcollet *et al.* 2002;Read *et al.* 2001b). As more evidence becomes available, the use of vaginal deliveries over elective CS may increase and indeed, guidelines are already changing with the offer of a vaginal delivery deemed acceptable in cases where women have low/undetectable viral load on stable HAART (Iribarren *et al.* 2001;The American College of Obstetricians and Gynecologists 2000; British HIV Association 2005). Decisions on the appropriate mode of delivery in infected pregnant women will be based on sequential measurements of viral load and a better understanding of changes of this marker in pregnancy may help facilitate these choices.

Prompt identification of undiagnosed HIV-infected women in pregnancy is crucial to the success of PMTCT. The few annual cases of vertically-acquired HIV infections in Western Europe largely due to late or non-identification of infection in the mother in pregnancy (Mayaux *et al.* 2003;European Collaborative Study 2005b). Although, for many reasons, the earlier an HIV-infected pregnant woman can be diagnosed the better, the finding of a faster time to viral suppression in ARV-naïve women initiating with NVP-based HAART highlights the importance of the choice of initial HAART regimen in women identified late in pregnancy.

Although guidelines state that pregnancy should not preclude use of optimal ARV regimens (DHHS Panel on Antiretroviral Guidelines for Adults and Adolescents 2006), in reality there are limited options. NFV has the most extensive pregnancy data on pharmacokinetics and safety of all PIs and is currently preferred within antenatal HAART, especially where there are no maternal indications for treatment (DHHS Panel on Antiretroviral Guidelines for Adults and Adolescents 2006). NFV was the overwhelming choice to accompany ZDV+3TC in ARV-naïve women to date here, but was less effective regarding virological suppression than NVP-containing HAART. In non-pregnant adults, NNRTI-containing HAART is recommended as a first-line regimen, preserving PI-containing HAART for later, with efavirenz as the preferred agent. As efavirenz is contraindicated in the first trimester of pregnancy, NVP-containing HAART has been increasingly used for ARV-naïve pregnant women in Europe, but this is now no longer recommended for women with relatively good immune functioning. The potential option of initiating efavirenz-based HAART in the second or third trimester, if contraception can be assured after delivery, has been suggested in current WHO recommendations for resource-limited countries (World Health Organisation 2006), but this approach is unlikely to be taken in Europe: among women in the ECS who initiated HAART during pregnancy in 2005, none did so with an efavirenz-based HAART regimen (unpublished data).

In the absence of clinical trials of HAART in pregnant women, these findings add to the evidence-base to assist therapeutic decision-making for ARV-naïve pregnant women. If clinicians comply with prescribing advice, one would expect the future group of ARV-naïve women starting on NVP-containing HAART to have lower CD4 counts than women here. These results strongly suggest that an ARV-naïve pregnant woman with

CD4 count <250 cells/mm^3 should be started on NVP-containing rather than NFV-containing HAART and highlight the urgent need for further research in pharmacokinetics, efficacy and safety of antiretroviral drugs in pregnancy. Additionally, once HAART becomes more readily available in low and middle-income countries, such data will also be important as a basis for comparisons with treatment responses in resource-poor settings where women are likely to be receiving HAART for the first time in pregnancy.

The use of HAART during pregnancy, particularly in the early stages when organogenesis is taking place, has raised concerns relating to the safety for the exposed fetus (Public Health Service Taskforce 2006b). The focus of safety issues in this thesis was neonatal outcomes only (i.e. prematurity and birth defects). Longer-term adverse effects of exposure to ARV in utero and/or neonatally have been documented by several groups including the ECS and include mitochondrial and haematological abnormalities (Blanche *et al.* 1999;Le Chenadec *et al.* 2003;Bunders *et al.* 2005).

Although further data are needed to confirm the findings here, based on these data there was no evidence of an association between HAART use and birth defects and an equivalent risk of prematurity between PI- and NNRTI-based HAART regimens was observed. Although both analyses included reasonably large sample sizes (3471 and 931, respectively), an increased risk of birth defects associated with treatment use in pregnancy cannot be ruled out, given that the study may not have had sufficient power to detect significant differences.

7.7 Data from the European Collaborative Study and their analysis: strengths and challenges

7.7.1 Data

The ECS clinical centres are the principal referral centres for HIV-1 infected pregnant women in their regions. Antenatal testing for HIV is routine in all centres and women enrolled in the ECS are likely to be representative of the HIV-infected pregnant clinic population in Europe who visited these centres for antenatal care. This is evidenced by the survey results from Section 2.3.1, which showed an enrolment rate of over 95% of HIV-infected pregnant women, with geographical relocation explaining most of the non-enrolments. As these data have been collected from a number of centres, characteristics of the centre-specific populations of infected pregnant women may vary and could have influenced these results. The ECS centres followed a standard protocol and the use of standardised data collection forms to collect important socio-demographic, obstetric and paediatric variables, details on treatment use, illicit drug use and immunological and virological markers of HIV disease, allowed for adjustment for these variables where appropriate in statistical analyses. However, most analyses did not explicitly adjust for centre-differences owing to the large numbers of centres within the ECS and differences in unmeasured confounders between centres could therefore not be accounted for which may have introduced some bias. For instance, clinician- and clinic-based preferences/policies for the timing and frequency of HIV RNA and CD4 count measurements or initial ART regimens may be confounded by unmeasured variables, which in turn could have an impact on all the results presented in this thesis. For example if some centres were more likely to take regular virological and immunological measurements than other centres, then this could lead to a bias in the results. A centre-based random effects term could be used to account for some of the unmeasured

confounders in LME models but was not feasible in the sub-group of women starting pregnancy on HAART (Chapter 5) as nested random effects cannot be incorporated into left-censored models at present (Thiébaut *et al.* 2005), and a nested random effects term for the sampling period was already included to account for differences in the storage time of samples in untreated women (Chapter 3). However, in an LME model with a random effect term for untreated women nested within centres, instead of the sampling period, the slope estimate for viral load in pregnancy was similar (estimate -0.016 $\log_{10}$ copies/ml per week; 95%CI -0.004, -0.029) as was the centre-associated random effect terms, suggesting that centre-associated variability in HIV RNA although significant, did not explain more variability than that associated with the sampling period.

The number of undetectable HIV RNA measurements (152; 31%) among the untreated group of women in the ECS was high (Chapter 3) and the extent to which the women in this study are representative of other untreated women with HIV infection during the same period is unknown. The prevalence of undetectable measurements reported in studies of untreated pregnant and non-pregnant adults varies. In a longitudinal study of 1256 non-pregnant adult women enrolled in the Women's Interagency HIV Study in the US, 48% and 18% of measurements taken prior to the initiation of ART were estimated to be <1466 and <80 copies/ml, respectively (Anastos *et al.* 2000). The proportion of undetectable HIV RNA measurements reported among 254 untreated pregnant women enrolled in the SEROGEST cohort (with a quantification limit of 400 copies/ml) was lower and varied from 16% in the second trimester of pregnancy to 15% at the time of delivery (Mayaux *et al.* 1997a). Eighty-percent of the viral load measurements included in the analysis of HIV RNA during pregnancy in Chapter 3 was from women enrolled before 1998, when viral load was not routinely collected within the study protocol. The

proportion of undetectable measurements among samples taken before 1998, where some had been stored for a longer duration, was considerably higher than among those from the later period (135/392; 34% vs. 17/98; 17%: χ^2=5.54, p=0.02). Accurate retrieval of HIV RNA has been reported from serum and plasma stored at -70°C for around 9 months (Coombs *et al.* 1993). In this study, early HIV RNA measurements were obtained from serum and plasma samples stored at -70°C for anything up to several years and the extent of the effect of a possible longer duration of storage on the accuracy of quantification of test results is not known (Coombs *et al.* 1993;Sabin *et al.* 2000). Although the analyses here did not directly control for heterogeneity of laboratory methods and storage times, all labs were university hospitals or tertiary referral centres with high levels of quality control. The analyses were therefore adjusted for HIV RNA sample and assay type where appropriate and a nested random effects structure used to account for unobserved differences between measurements taken in the two time periods, which should reduce some of the bias in the estimates from these models. Nevertheless, the finding of a large number of undetectable measurements in this study could feasibly represent the degradation of samples from several years' storage and this should be taken into account when interpreting the results from the analysis of untreated women. Additionally, there was no record of any ART use during pregnancy or in previous pregnancies among the 15 women in this analysis who had their second child in the ECS, and thorough data checks were carried out (see Section 2.3.1). However, it could be possible that the ART start dates for these women or any treatment received pre-pregnancy were not accurately recorded and that some of the decreases observed in pregnancy may be treatment-related. Even with these possible limitations of the data, the proven success of ART means that it is unlikely that additional information on untreated pregnant women will become available and

measurements on this group of women offered a unique opportunity to study the natural dynamics of HIV RNA in pregnancy and add to the limited knowledge on viral load patterns in pregnancy.

In the analysis of HIV RNA and CD4 counts among women becoming pregnant on HAART (Chapter 5), information on the duration of HAART pre-pregnancy was not available for all women and therefore the resulting parameter estimates may be confounded by time on treatment. A sub-analysis of 109 women with available information on timing of treatment revealed mostly negligible changes in the parameter estimates for the intercept and slopes of viral load and CD4 count over pregnancy. The only exception was the estimate of the second slope of CD4 count in pregnancy which was of similar magnitude but had a 95% CI which was no longer bounded by zero; this suggests that any confounding effect of treatment duration on the results here was negligible in the viral load model, but may have resulted in a slope estimate biased towards the null for the CD4 model.

The gold standard for comparison of the effectiveness of different treatments is a randomised controlled trial (RCT) and the comparison of HAART regimens in pregnancy on the time taken to achieve undetectable viral load by delivery is limited by the observational nature of the data (Phillips *et al.* 1999). RCTs, in which randomisation can be expected to remove all confounding factors, is the only design that can guarantee unbiased, reliable estimates of treatment effects (Chalmers *et al.* 1983). However, there are also a number of well-documented problems associated with RCTs (Sabin and Phillips 2001). The relevance of findings from an RCT in a group of patients who have consented to participate and who have been selected based on restricted immunological

and virological criteria may not reflect clinical practice, and their relevance to the general patient population is debatable (Sabin and Phillips 2001;Sabin 2004;Munoz *et al.* 2000;Phillips *et al.* 1999). In the absence of data from RCTs and other observational studies, the findings in this thesis on the effects of choice of initial HAART regimen on achieving a timely undetectable viral load in pregnancy are important. The use of analytical methods which allowed for interval-censoring (Lindsey and Ryan 1998), adjusted for timing of initiation in pregnancy and minimised confounding by including treatment-naïve women should improve the estimated treatment effects. A disadvantage of these data is their limited contemporary relevance where therapeutic practices have changed over time. A case-in-point is the NVP prescribing changes following the association with hepatotoxicity in women with moderate to high CD4 counts (Public Health Service Taskforce 2006b). Although specific data were not available on ALT/AST levels and drug-related toxicities, all ARV-naïve women who started NVP-based HAART in this analysis continued with this treatment to at least delivery, suggesting that any toxicities arising were minor and/or not necessitating treatment modification.

The results in the time to undetectable viral load analysis depend on the assumption that women receiving NVP or PI-based HAART with the same values of measured prognostic factors were similar. Additionally, the results from the analysis of viral load and CD4 counts among women becoming pregnant on HAART assume that all confounders were adjusted for so that any changes seen in pregnancy can be attributed to measured covariates. However, factors potentially influencing response to HAART could not be accounted for such as adherence, biological differences in drug activity arising from variations in body weight and pharmacokinetics between groups, HIV sub-

types and other genetic factors (Stebbing and Bower 2006;Lowe *et al.* 2004;Mirochnick and Capparelli 2004). Residual confounding by these or other factors is unlikely to be entirely eliminated from observational studies and the results must be interpreted in this light.

In a recent review paper, the most common reason given for failure to achieve or sustain expected treatment benefits was non-adherence (Conway 2007). Therapeutic success, including virological and immunologic responses, has been demonstrated across PI- and NVP-containing HAART regimens for patients who maintain good levels of adherence (generally >95%) (Conway 2007;Moore *et al.* 2005;Wood *et al.* 2004). This has been corroborated in studies of directly observed therapy and directly administered therapy where levels of adherence are near 100% (Lucas *et al.* 2006;Conway 2007). Pregnancy is associated with significant physiologic changes, and nausea and vomiting are common. This in itself could lead to decreased drug levels in the first trimester of pregnancy and any nausea and vomiting may also increase the patient's reluctance to maintain a pill regimen which is associated with a perceived additional side-effect (Hammer *et al.* 2006;Conway 2007). There is also the possibility that pregnant women may be less likely to adhere in pregnancy due to concerns over the effect these treatment regimens may have on their unborn child. However, a recent study of adherence from Brazil based on 72 pregnant women and 79 non-pregnant women found that pregnant women had a more than three times odds of adhering than non-pregnant women (Vaz *et al.* 2007). Given these results, it is feasible that the women who were receiving HAART at conception (Chapter 5) may have increased their adherence in the first trimester of pregnancy once they were aware of their pregnancy, and may help explain why a two-slope pattern was necessary to describe viral load and CD4 count

changes in pregnancy. However, results from studies examining adherence for women who were on ART before pregnancy are limited and conflicting. The Brazilian study mentioned above found a similar rate of adherence among women who were on ART before pregnancy compared with those who began ART during pregnancy (Vaz *et al.* 2007), while a New York cohort of 2714 HIV-infected women found that women on ART before pregnancy were more likely to adhere (Laine *et al.* 2000). Whether adherence was a confounder in this study could not be determined, but this uncertainty should be taken into account when making inferences based on these results.

Unobserved differences in the rate of adherence, pill burden, and associated side-effects in the receipt of PI- and NVP-based HAART may also bias the results from the time to undetectable viral load analysis (Chapter 4). However a meta-analysis of virologic outcomes in relation to pill burden in studies of triple-drug HAART in treatment-naive HIV-infected persons found that pill count was not a consistent predictor of virologic suppression (Bartlett *et al.* 2006), suggesting that any differences in pill burden may have only had a limited influence on the observed results here. Nevertheless, without information on pharmacokinetics of drugs in pregnancy in the ECS, it was not possible to confirm findings from studies suggesting that NFV drug levels in the third trimester are sub-therapeutic (Villani *et al.* 2006;Nellen *et al.* 2004;van Heeswijk *et al.* 2004), and it is possible that differences in the rate of adherence between treatment regimen groups were a confounding factor. Additionally, data on seroconversion was not available, and the data of first HIV diagnosis is unlikely to be a good proxy for this. Studies of viral load trends in mostly male seroconverters found that viral load decreased in the first few years from seroconversion, after which they gradually increased (Lyles *et al.* 1999;Sabin *et al.* 2000). It is not possible to verify whether West

African women in the time to undetectable viral load analysis had seroconverted earlier than other women, but may offer an additional explanation to why this group of women were seen to respond faster to treatment than Non-African women (Chapter 4).

7.7.2 Statistical methods

The analysis of HIV data in pregnancy was complicated by several factors, which required the use of complex, non-standard statistical methods.

The monitoring of HIV markers in pregnant women in the ECS resulted in repeated measurements of immunological and virological markers and were analysed using LME models which took into account the correlation between measurements on each woman (Chapter 3 and 5). In the analysis of viral load measurements in untreated pregnant women, the use of a nested random effects model was justified by the considerable and significant variation between the two periods of HIV RNA data collection, but precluded the use of complex methods to take into account the left-censoring resulting from undetectable measurements (Chapter 3). In an attempt to remove the heteroscedasticity in residuals resulting from use of the simple method of imputing midpoints, a power function was incorporated into the model to account for this variance structure, which led to an improvement in the fit of the model. The lack of homoscedasticity among residuals from undetectable values seems unavoidable when relatively simple methods such as imputation of midpoints are used. Although this is unlikely to invalidate the estimates of the models' parameters it may impact the associated standard errors and 95% confidence intervals, and the use of bootstrap methods to obtain confidence intervals for LME models where censoring is high, may be more appropriate (Ware 1985).

A sensitivity analysis was carried out to assess the impact of imputing midpoints for undetectable measurements in the analysis of viral load in untreated women (Chapter 3). Using the conditional likelihood method to account for undetectable measurements and accounting for sampling period through a covariate term, a similar slope in pregnancy was revealed but with a considerably larger estimate for the mean HIV RNA at delivery intercept. This change in the intercept could be a direct result of either not accounting for the variability in sampling periods appropriately or using a more appropriate statistical method for the analysis of undetectable measurements (Jacqmin-Gadda *et al.* 2000). These results suggested that although the slope estimate in this analysis was invariant to the imputation method, the mean estimate for the intercept was not and therefore caution should to be taken when interpreting the estimate for mean HIV RNA at delivery.

The decision to use a piecewise linear LME model for viral load and CD4 measurements over pregnancy in treated women was based on the underlying pattern of the markers, and consideration of both minimisation of the AIC and ease of interpretation (Chapter 5). A limitation of this approach is that the chosen functional form may not have been fully representative of the markers on the introduction of explanatory variables, in particular with respect to the time of change-point. In addition, the median number of CD4 and HIV RNA measurements among this sample of treated pregnant women was only three limiting any inferences that can be made on the underlying population trends in markers.

The conditional likelihood method resulted in improved properties of the parameter estimates for viral load in women becoming pregnant while receiving HAART, although there was some evidence of non-normality in the tails of the residuals (Chapter 5). However, after adjustment for CD4 count this was less evident.

The statistical method used to account for left-censoring among treated women was based on the assumption that the left-censored values are part of a Gaussian distribution, which estimates these viral load values by completing this distribution using the likelihood method (Thiébaut and Jacqmin-Gadda 2004). An asymmetrical distribution such as the gamma distribution may be more suitable for undetectable viral load values, however at present the approach using the NLMIXED procedure in SAS for left-censored values is not able to incorporate any other distributions and this modification was not investigated further.

A bivariate model was used to account for and to estimate the correlation between the markers in pregnancy but did not result in an improved model fit, while a bivariate model with the conditional likelihood approach could not be fitted (Chapter 5). A possible explanation for this is the small variation around the CD4 slope and intercept. The bivariate approach followed here takes into account the correlation between the two markers through the covariance matrix of their random effects, and insufficient variability around the random terms is likely to result in estimated covariances (correlations) close to zero, leading to unreliable and unstable parameter estimates and their standard errors (Everitt 1987;Goldstein 1989;SAS Institute Inc. 2004). Bivariate LME models with left-censoring have also been used to examine the changes in viral load and CD4 from baseline in non-pregnant adults from the APROCO Cohort and the

CASCADE collaboration after initiation of treatment (Thiébaut *et al.* 2003;Thiébaut *et al.* 2006), using the CENSAD routine which requires the computation of a multiple integral over the number of censored measures per subject (Jacqmin-Gadda *et al.* 2000). The APROCO analysis included 929 patients with 4513 and 4491 CD4 count and viral load measurements followed for up to 12 months after initiation (Thiébaut *et al.* 2003), and the CASCADE analysis 943 patients with over 7000 viral load and CD4 measurements and up to 24 months follow up after initiation (Thiébaut *et al.* 2006). No appreciable changes in parameter estimates from the use of a bivariate model over the univariate model were reported in these studies, but the correlations between the intercepts and slopes for both markers were adequately estimated. Several differences between these studies and the ECS analysis may explain why the random effects covariance could be properly estimated in these studies, but not in the ECS. The period of follow-up in these studies was considerably longer than in the ECS analysis which was restricted to the period of pregnancy and as a result the median number of measurements was only 3 in the ECS analysis, whereas the median number of measurements for viral load and CD4 counts was 5 in the APROCO study and 4 in the CASCADE study. Additionally, the statistical approach used in these two studies used the CENSAD routine, which uses a Marquadt algorithm to maximise the likelihood, a method that has been shown to be preferable when the number of random effects are large (>4 random effect terms) (Thiébaut and Jacqmin-Gadda 2004). However, in the ECS analysis the number of random effects was limited to 4 or less and the NLMIXED procedure has been shown to be reliable up to 4 random effects so the difference in statistical method is an unlikely explanation for why the random effects covariance could not be estimated for the ECS data (Thiébaut and Jacqmin-Gadda 2004). Although the possibility remains that the integration algorithm used in the CENSAD routine may

have been more suitable for these data in estimating the random effects covariance matrix, the small amount of variation around the intercept and slopes for the women in these studies is a more likely explanation.

The bivariate random effects model, taking into account left-censoring with the conditional likelihood method is the preferable approach to the joint modelling of CD4 and viral load as it allows for direct interpretation of the relationship between the two markers without assuming the dependence of one marker in relation to the other (Thiébaut *et al.* 2002). However as this model could not be estimated LME models of each marker as a function of the other marker were fitted instead. Given that HIV RNA directly targets CD4 cells leading to their decline, the LME model MODCD4-2, which modelled CD4 as a function of viral load may be biologically more credible (Chapter 5). However, during the period of pregnancy, the main focus of PMTCT is to ensure women have as low viral loads as possible throughout pregnancy and by the time of delivery (Mofenson *et al.* 1999;European Collaborative Study 1999;European Collaborative Study 2005b), and for this reason the LME model for viral load, MODVL-CL2, is the preferred model for informing clinical practice.

In the analysis of treated women, the repeated measures nature of the data was taken into account by assuming an unstructured covariance matrix for the random intercept and slope terms for all LME models (Chapter 5). The use of specific correlation structures to model the dependence among observations (through their within-group errors) was not feasible here as these methods require that the data are observed sequentially over time, i.e. at equally spaced time points (Pinheiro and Bates 2000). Compound symmetry and autoregressive LME models therefore could not be

investigated. Further, the addition of stochastic terms to the LME models to allow for the correlation between measurements, using for example an Integrated Ornstein Uhlenbeck (IOU) or Brownian motion term, was not feasible (Boscardin *et al.* 1998). Given sequentially observed data, Boscardin et al. also discuss the possibility of the joint modelling of two parameters by using a bivariate IOU process (Boscardin *et al.* 1998). This bivariate stochastic model allows for the stochastic process in one marker to be influenced by the previous derivatives of the processes of both markers, while the process of the other marker is influenced only by its own process. This may be a suitable approach in the joint modelling of HIV RNA viral load and CD4, where the lag between measurements and the influence of previous measurements of each marker on subsequent CD4 and viral load measurements may be different for each marker. The bivariate stochastic model could allow for asymmetry in how one marker affects the other and also the lag between these measurements; whether this approach is suitable for the joint analysis of HIV RNA viral load and CD4 in pregnancy needs to be determined.

A limitation of the analysis of treated women in pregnancy (Chapter 5) was that it was carried out according to the intention-to-treat (or continue) principle and therefore did not take into account any treatment modifications or interruptions; the results therefore could be biased towards determining no effect (Pezzotti *et al.* 2001).

The LDM proved to be a useful tool in assessing the assumption of independence between residuals from the bivariate LME model (Chapter 5).

As expected in an observational cohort study, treatment allocation was not randomised and viral load measurements occurred at various times after initiation of treatment depending on the frequency of clinic visits in pregnancy (Chapter 4). The non-standard measurement schedule and the interval censoring nature of the data required a parametric survival model to assess treatment differences (Griffin *et al.* 2006;Lindsey and Ryan 1998;Klein and Moeschberger 1997). Alternative parametric distributions, such as the log-logistic distribution or the piecewise exponential (Farrington 1996) have also been shown to be suitable for assessing treatment outcomes for data with this censoring structure (Griffin *et al.* 2006). However, investigation of the hazard rate here (Chapter 4) showed the Weibull model to be appropriate in modelling the hazard rates for the two treatment regimens, but with the advantage that the model estimates had a simpler interpretation through representation as an AFT or proportional hazards model (Klein and Moeschberger 1997).

The finding of an association between initiation of HAART in the third trimester and more rapid achievement of undetectable virus was a surprising finding. It is possible that this effect represents uncontrolled confounding or that this is an artefact of the time in the study for each woman related to the timing of viral load measurements that was not fully accounted for. The Weibull model used here also assumed that the viral load measurement times were independent of a woman's treatment outcome, but it is feasible that the number of clinic visits for each woman was related to their current viral load, which will also have been dependent on the effectiveness of the treatment (Manavi *et al.* 2007;Farrington and Gay 1999). Farrington et al. have proposed a statistical method to model interval-censored survival data with informative examination times, which involves modelling the hazard function by taking account of individual frailties related

to the frequency of clinic measurements; the extent to which this would impact on estimated treatment differences in this setting, is not known (Farrington and Gay 1999).

Adjustment for the propensity score was used to create a quasi-randomised trial setting in the aim of reducing some of the bias associated with non-randomised treatment allocation but had a negligible effect on the treatment effect estimate (Chapter 4) (D'Agostino, Jr. 1998). However, propensity score methods assume that any differences in treatment allocation can be measured by the observed covariates, and biased treatment effect estimates may still result if this is not the case, which may explain why the propensity score correction made little difference in this analysis. Stürmer et al. have proposed a method called propensity score calibration, which addresses confounding by variables unobserved in the study, by using variables observed in a validation study (Sturmer *et al.* 2005). In the absence of an appropriate validation study for HIV-infected pregnant women, this is not a feasible approach at present.

An additional explanation for the negligible change in the treatment effect estimate after adjustment for the propensity score is that the logistic regression model, which predicts the propensity scores, may have been misspecified. Drake et al. carried out a simulation study of 1000 samples of size 100 to compare the propensity score model (stratification by quintiles of the propensity score) and prognostic model (adjusting for covariates only) in estimating treatment effects from observational studies (Drake 1993). After comparison of the resulting estimates, the author concluded that misspecification of the propensity score model resulted in smaller biases than misspecification of the prognostic model, however the bias associated with omitting a confounder was similar between both models. Given these results, it is unlikely that misspecification of the

propensity score will have resulted in any treatment effect bias in the study presented here as all observed variables which were confounders were incorporated into the Weibull model.

There are currently few data on the ability of the propensity score to reduce treatment allocation bias in survival models. A recent study by Austin et al. used Monte Carlo simulations to examine the potential bias of different propensity score methods for estimating treatment effects through logistic, Poisson and Cox proportional hazards regression models (Austin *et al.* 2007). They found that when either the odds ratio or the hazards ratio was used as the measure of treatment effect and when there was a true underlying treatment effect, conditioning on the propensity score resulted in biased estimation of the true treatment effect, towards that of no effect. This held true regardless of whether the scores were incorporated into the model through stratification, matching or covariate adjustment. However, they also found that for time-to-event outcomes, including only the true confounders in the regression model also resulted in a treatment effect bias; the model which adjusted for all variables, i.e. both those related to treatment outcome and those that are true confounders, resulted in no bias in the HR for the treatment effect (Austin *et al.* 2007). The propensity score model in the analysis of treatment effects on time to undetectable viral load presented here included only variables which were related to treatment effect, even if they were not related to the outcome of interest. This ensured that although the final Weibull model was not over parameterised, treatment assignment was balanced among women, given the observed covariates. However, given that the results from this study showed no changes in the treatment effect estimate, and based on the results from the study by Austin et al., propensity score adjustment seems to add little in terms of reducing bias in treatment

effect estimates, and adjusting for confounders and variables which are related to the outcome appears to be the most appropriate approach.

Assessing the outcome of premature delivery in this cohort was complicated by the considerable use of elective CS (Chapter 6). These deliveries were viewed as right-censored outcomes, with a Cox-proportional hazards model used to assess the affect of type of HAART on the rate of delivery.

An important assumption made in statistical analysis is that any data which are missing are missing at random. In the analysis of prematurity between women receiving NNRTI-based vs. PI-based HAART, there was some evidence of dependence between missing CD4 count and timing of treatment information and the type of HAART being received (Chapter 6). Therefore, this missing information and the mechanism of missingness was accounted for by weighting observations in the regression based on their likelihood of being incomplete, using inverse probability weights predicted from a propensity score model. The LME models used in the analysis of HIV RNA levels and CD4 cell counts in untreated and treated women also assume that the responses missing at any time point are missing at random. This assumption is important as although pregnant women in the ECS will have had a similar number of clinic visits as part of their routine antenatal care, not all women will have had an HIV RNA or CD4 cell measurement taken. However, if the measurements which were taken were not representative of the woman's true trajectory or response to therapy, then a bias could be introduced. In the analysis of untreated women, HIV RNA levels among women with one measurement only were lower at delivery than women who had at least two measurements, and their inclusion in the model may lead to a possible bias. However, in

a sensitivity analysis excluding women with only one measurement, the only change was in the estimate of mean HIV RNA at delivery, while the slope estimate was similar.

7.8 Recommendations and future research

There are several methodological aspects of this work which could be explored further but which was outside the scope of this thesis. Most HIV RNA assays used today are likely to be ultrasensitive, and the need for complicated methods to account for left-censoring in datasets where detection limits of assays are predominantly ≤50 copies/ml may not be as important as in the ECS datasets investigated here (Gray *et al.* 2004). Nevertheless, a clustering of measurements with the same undetectable value can lead to lack of homoscedasticity as observed here, and simple but effective methods to account for this need to be developed.

A limitation of the analyses of viral load and CD4 markers over pregnancy was that no preconception or postpartum data were available. Although, inferences on the effect of pregnancy are possible based on these data, prospective studies of women before, during and after pregnancy are needed to determine whether or not the changes observed here are due to pregnancy (either directly or indirectly).

There is an urgent need for further research in pharmacokinetics, efficacy and safety of antiretroviral drugs in pregnancy. Boosted PI regimens appear to offer superior virological suppression in ARV-naïve adults compared to PI alone (Bartlett *et al.* 2006) and lopinavir/ritonavir is identified as a preferred PI for initial HAART in pregnancy in current US guidelines, albeit with limited pharmacokinetic and safety data (DHHS Panel on Antiretroviral Guidelines for Adults and Adolescents 2006;Stek *et al.* 2006).

While only 17 women received boosted PIs in the ECS in 2005 (unpublished data), as more information becomes available, boosted PIs in initial HAART regimens in pregnancy may become increasingly common and the question of its equivalence to other HAART regimens warrants investigation. The UK CHIC study has reported a poor immunological and virological status in a small group of patients who have experienced failure of all three classes of ARVs and suggested the possibility of treatment exhaustion for these individuals (Sabin *et al.* 2005). Given the trend in increasing subsequent live births in infected women in Europe (European Collaborative Study 2005a) and the limited options for ARV treatment in pregnancy, research into newer agents such as integrase and fusion inhibitors with regards to both their safety and tolerability in pregnancy and their effectiveness for PMTCT need consideration.

Reference List

Abberger, K. (2002) Exploring local dependence. *CoFE Discussion Paper 02-14*, Center of Finance and Econometrics, University of Konstanz.

AIDSinfo . AIDSinfo drug database. http://www.hivatis.org/. Accessed February 2007.

Akaike, H. (1974) A new look at the statistical model identification. *IEEE Trans.Automat.Contr.*, **AC-19,** 716-23.

Anastos, K., Gange, S. J., Lau, B., Weiser, B., Detels, R., Giorgi, J. V., Margolick, J. B., Cohen, M., Phair, J., Melnick, S., Rinaldo, C. R., Kovacs, A., Levine, A., Landesman, S., Young, M., Munoz, A. and Greenblatt, R. M. (2000) Association of race and gender with HIV-1 RNA levels and immunologic progression. *J.Acquir.Immune Defic.Syndr.*, **24,** 218-226.

Antiretroviral Pregnancy Registry Steering Committee (2004). Antiretroviral Pregnancy Registry Interim Report for 1st January 1989 to 31st January 2003.

Anzala, A. O., Simonsen, J. N., Kimani, J., Ball, T. B., Nagelkerke, N. J., Rutherford, J., Ngugi, E. N., Bwayo, J. J. and Plummer, F. A. (2000) Acute sexually transmitted infections increase human immunodeficiency virus type 1 plasma viremia, increase plasma type 2 cytokines, and decrease CD4 cell counts. *J.Infect.Dis.*, **182,** 459-466.

Austin, P. C., Grootendorst, P., Normand, S. L. and Anderson, G. M. (2007) Conditioning on the propensity score can result in biased estimation of common measures of treatment effect: a Monte Carlo study. *Stat.Med.*, **26,** 754-768.

Austin, P. C., Mamdani, M. M., Stukel, T. A., Anderson, G. M. and Tu, J. V. (2005) The use of the propensity score for estimating treatment effects: administrative versus clinical data. *Stat.Med.*, **24,** 1563-1578.

Barre-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W. and Montagnier, L. (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*, **220,** 868-871.

Bartlett, J. A., Fath, M. J., Demasi, R., Hermes, A., Quinn, J., Mondou, E. and Rousseau, F. (2006) An updated systematic overview of triple combination therapy in antiretroviral-naive HIV-infected adults. *AIDS*, **20,** 2051-2064.

Baylor, M. S. and Johann-Liang, R. (2004) Hepatotoxicity associated with nevirapine use. *J.Acquir.Immune Defic.Syndr.*, **35,** 538-539.

Bersoff-Matcha, S. J., Miller, W. C., Aberg, J. A., van Der, H. C., Hamrick Jr, H. J., Powderly, W. G. and Mundy, L. M. (2001) Sex differences in nevirapine rash. *Clin.Infect.Dis.*, **32,** 124-129.

Bessinger, R., Clark, R., Kissinger, P., Rice, J. and Coughlin, S. (1998) Pregnancy is not associated with the progression of HIV disease in women attending an HIV outpatient program. *Am.J.Epidemiol.*, **147,** 434-440.

Blanche, S., Tardieu, M., Rustin, P., Slama, A., Barret, B., Firtion, G., Ciraru-Vigneron, N., Lacroix, C., Rouzioux, C., Mandelbrot, L., Desguerre, I., Rotig, A., Mayaux, M. J. and Delfraissy, J. F. (1999) Persistent mitochondrial dysfunction and perinatal exposure to antiretroviral nucleoside analogues. *Lancet*, **354,** 1084-1089.

Bocket, L., Cheret, A., Deuffic-Burban, S., Choisy, P., Gerard, Y., de, l. T., X, Viget, N., Ajana, F., Goffard, A., Barin, F., Mouton, Y. and Yazdanpanah, Y. (2005) Impact of

human immunodeficiency virus type 1 subtype on first-line antiretroviral therapy effectiveness. *Antivir.Ther.*, **10,** 247-254.

Boer, K., Nellen, J. F., Patel, D., Timmermans, S., Tempelman, C., Wibaut, M., Sluman, M. A., van der Ende, M. E. and Godfried, M. H. (2007) The AmRo study: pregnancy outcome in HIV-1-infected women under effective highly active antiretroviral therapy and a policy of vaginal delivery. *BJOG.*, **114,** 148-155.

Boscardin, W. J., Taylor, J. M. and Law, N. (1998) Longitudinal models for AIDS marker data. *Stat.Methods Med.Res.*, **7,** 13-27.

Brettle, R. P., Raab, G. M., Ross, A., Fielding, K. L., Gore, S. M. and Bird, A. G. (1995) Hiv-Infection in Women - Immunological Markers and the Influence of Pregnancy. *AIDS*, **9,** 1177-1184.

British HIV Association. Guidelines for the management of HIV infection in pregnant women and the prevention of mother-to-child transmission of HIV. http://www.bhiva.org/guidelines/2005/pregnancy/Pregnancy2005.doc Accessed March 2007.

Brockmeyer, N. (1999) German-Austrian Guidelines for HIV-therapy during pregnancy--status: May/June 1998--common statement of the Deutsche AIDS-Gesellschaft (DAIG) and the Osterreichische AIDS-Gesellschaft (OAG). *Eur.J.Med.Res.*, **4,** 35-42.

Brumme, Z. L. and Harrigan, P. R. (2006) The impact of human genetic variation on HIV disease in the era of HAART. *AIDS Rev.*, **8,** 78-87.

Bryant, A. S., Leighty, R. M., Shen, X., Read, J. S., Brouwers, P., Turpin, D. B., Larussa, P. S., Pacheco-Acosta, E., Paul, M. E., Vajaranant, M. and Tuomala, R. E. (2007) Predictors of Repeat Pregnancy Among HIV-1-Infected Women. *J.Acquir.Immune Defic.Syndr.*, **44,** 87-92.

Buckley, J. and James, I. (1979) Linear-Regression with Censored Data. *Biometrika*, **66,** 429-436.

Bunders, M., Lugada, E., Mermin, J., Downing, R., Were, W., Thorne, C. and Newell, M. L. (2006) Within and between race differences in lymphocyte, CD4+, CD8+ and neutrophil levels in HIV-uninfected children with or without HIV exposure in Europe and Uganda. *Ann.Trop.Paediatr.*, **26,** 169-179.

Bunders, M., Thorne, C. and Newell, M. L. (2005) Maternal and infant factors and lymphocyte, CD4 and CD8 cell counts in uninfected children of HIV-1-infected mothers. *AIDS*, **19,** 1071-1079.

Burns, D. N., Landesman, S., Minkoff, H., Wright, D. J., Waters, D., Mitchell, R. M., Rubinstein, A., Willoughby, A. and Goedert, J. J. (1998) The influence of pregnancy on human immunodeficiency virus type 1 infection: antepartum and postpartum changes in human immunodeficiency virus type 1 viral load. *Am.J.Obstet.Gynecol.*, **178,** 355-359.

Burns, D. N., Nourjah, P., Minkoff, H., Korelitz, J., Biggar, R. J., Landesman, S., Rubinstein, A., Wright, D. and Nugent, R. P. (1996) Changes in CD4+ and CD8+ cell levels during pregnancy and post partum in women seropositive and seronegative for human immunodeficiency virus-1. *Am.J.Obstet.Gynecol.*, **174,** 1461-1468.

Cao, Y., Krogstad, P., Korber, B. T., Koup, R. A., Muldoon, M., Macken, C., Song, J. L., Jin, Z., Zhao, J. Q., Clapp, S., Chen, I. S., Ho, D. D. and Ammann, A. J. (1997) Maternal HIV-1 viral load and vertical transmission of infection: the Ariel Project for the prevention of HIV transmission from mother to infant. *Nat.Med.*, **3,** 549-552.

Castilla, J. A., Rueda, R., Vargas, M. L., Gonzalez-Gomez, F. and Garcia-Olivares, E. (1989) Decreased levels of circulating CD4+ T lymphocytes during normal human pregnancy. *J.Reprod.Immunol.*, **15,** 103-111.

Centers for Disease Control (1981) Pneumocystis pneumonia--Los Angeles. *MMWR Morb.Mortal.Wkly.Rep.*, **30,** 250-252.

Centers for Disease Control (1982) Unexplained immunodeficiency and opportunistic infections in infants--New York, New Jersey, California. *MMWR Morb.Mortal.Wkly.Rep.*, **31,** 665-667.

Centers for Disease Control (1992) 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *MMWR Recomm.Rep.*, **41,** 1-19.

Chalmers, T. C., Celano, P., Sacks, H. S. and Smith, H., Jr. (1983) Bias in treatment assignment in controlled clinical trials. *N.Engl.J.Med.*, **309,** 1358-1361.

Chirenda, J. (1999) Low CD4 count in HIV negative malaria cases and normal CD4 count in HIV positive and malaria negative patients. *Cent.Afr.J.Med.*, **45,** 248.

Chou, R., Fu, R., Huffman, L. H. and Korthuis, P. T. (2006) Initial highly-active antiretroviral therapy with a protease inhibitor versus a non-nucleoside reverse transcriptase inhibitor: discrepancies between direct and indirect meta-analyses. *Lancet,* **368,** 1503-1515.

Coll O, Suy A and Martinez E . Increased risk of pre-eclampsia and fetal death in HIV-infected pregnant women receiving highly-active antiretroviral therapy. Proceedings and Abstracts of the 11th Conference on Retroviruses and Opportunistic Infections , San Francisco, CA, USA, 8-11 February 2004 (Abstract 921).

Collaborative Group on AIDS Incubation and HIV Survival including the CASCADE EU Concerted Action (2000) Time from HIV-1 seroconversion to AIDS and death before widespread use of highly-active antiretroviral therapy: a collaborative re-analysis. *Lancet,* **355,** 1131-1137.

Collett, D. (1994) *Modelling Survival Data in Medical Research.* Chapman & Hall, London.

Collier, A. C., Coombs, R. W., Schoenfeld, D. A., Bassett, R. L., Timpone, J., Baruch, A., Jones, M., Facey, K., Whitacre, C., McAuliffe, V. J., Friedman, H. M., Merigan, T. C., Reichman, R. C., Hooper, C. and Corey, L. (1996) Treatment of human immunodeficiency virus infection with saquinavir, zidovudine, and zalcitabine. AIDS Clinical Trials Group. *N.Engl.J.Med.*, **334,** 1011-1017.

Connor, E. M., Sperling, R. S., Gelber, R., Kiselev, P., Scott, G., O'Sullivan, M. J., VanDyke, R., Bey, M., Shearer, W., Jacobson, R. L. and . (1994) Reduction of maternal-infant transmission of human immunodeficiency virus type 1 with zidovudine treatment. Pediatric AIDS Clinical Trials Group Protocol 076 Study Group. *N.Engl.J.Med.*, **331,** 1173-1180.

Conway, B. (2007) The role of adherence to antiretroviral therapy in the management of HIV infection. *J.Acquir.Immune Defic.Syndr.*, **45 Suppl 1,** S14-S18.

Coombs, R. W., Henrard, D. R., Mehaffey, W. F., Gibson, J., Eggert, E., Quinn, T. C. and Phillips, J. (1993) Cell-free plasma human immunodeficiency virus type 1 titer assessed by culture and immunocapture-reverse transcription-polymerase chain reaction. *J.Clin.Microbiol.*, **31,** 1980-1986.

Cooper, E. R., Charurat, M., Mofenson, L., Hanson, I. C., Pitt, J., Diaz, C., Hayani, K., Handelsman, E., Smeriglio, V., Hoff, R. and Blattner, W. (2002) Combination antiretroviral strategies for the treatment of pregnant HIV-1-infected women and prevention of perinatal HIV-1 transmission. *J.Acquir.Immune Defic.Syndr.*, **29,** 484-494.

Cooper, E. R., Nugent, R. P., Diaz, C., Pitt, J., Hanson, C., Kalish, L. A., Mendez, H., Zorrilla, C., Hershow, R., Moye, J., Smeriglio, V. and Fowler, M. G. (1996) After AIDS clinical trial 076: the changing pattern of zidovudine use during pregnancy, and the subsequent reduction in the vertical transmission of human immunodeficiency virus in a

cohort of infected women and their infants. Women and Infants Transmission Study Group. *J.Infect.Dis.*, **174,** 1207-1211.

Cotter, A. M., Garcia, A. G., Duthely, M. L., Luke, B. and O'Sullivan, M. J. (2006) Is antiretroviral therapy during pregnancy associated with an increased risk of preterm delivery, low birth weight, or stillbirth? *J.Infect.Dis.*, **193,** 1195-1201.

Covington, D. L., Conner, S. D., Doi, P. A., Swinson, J. and Daniels, E. M. (2004) Risk of birth defects associated with nelfinavir exposure during pregnancy. *Obstet.Gynecol.*, **103,** 1181-1189.

Cox DR and Oakes D (1984) *Analysis of survival data*. Chapman and Hall.

Cox, C., Chu, H., Schneider, M. F. and Munoz, A. (2007) Parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution. *Stat.Med.* **26,** 4352-4374.

Cox, D. R., Atkinson, A. C., Box, G. E. P., Darroch, J. N., Spjotvoll, E. and Wahrendorf, J. (1984) Interaction. *International Statistical Review*, **52,** 1-31.

Cunningham, C. K., Balasubramanian, R., Delke, I., Maupin, R., Mofenson, L., Dorenbaum, A., Sullivan, J. L., Gonzalez-Garcia, A., Thorpe, E., Rathore, M. and Gelber, R. D. (2004) The impact of race/ethnicity on mother-to-child HIV transmission in the United States in Pediatric AIDS Clinical Trials Group Protocol 316. *J.Acquir.Immune Defic.Syndr.*, **36,** 800-807.

D'Agostino, R. B., Jr. (1998) Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat.Med.*, **17,** 2265-2281.

Davison, A. C. and Hinkley, D. V. (1997) *Bootstrap Methods and their Applications, Chapter 5.* Cambridge University Press.

De Santis, M., Carducci, B., De Santis, L., Cavaliere, A. F. and Straface, G. (2002) Periconceptional exposure to efavirenz and neural tube defects. *Arch.Intern.Med.*, **162,** 355.

Delta Study (1996) Delta: a randomised double-blind controlled trial comparing combinations of zidovudine plus didanosine or zalcitabine with zidovudine alone in HIV-infected individuals. Delta Coordinating Committee. *Lancet*, **348,** 283-291.

DHHS Panel on Antiretroviral Guidelines for Adults and Adolescents . Guidelines for the Use of Antiretroviral Agents in HIV-1-Infected Adults and Adolescents. http://aidsinfo.nih.gov/ Accessed October 2006.

Dorenbaum, A., Cunningham, C. K., Gelber, R. D., Culnane, M., Mofenson, L., Britto, P., Rekacewicz, C., Newell, M. L., Delfraissy, J. F., Cunningham-Schrader, B., Mirochnick, M. and Sullivan, J. L. (2002) Two-dose intrapartum/newborn nevirapine and standard antiretroviral therapy to reduce perinatal HIV transmission: a randomized trial. *JAMA*, **288,** 189-198.

Downs, A. M., Heisterkamp, S. H., Rava, L., Houweling, H., Jager, J. C. and Hamers, F. F. (2000) Back-calculation by birth cohort, incorporating age- specific disease progression, pre-AIDS mortality and change in European AIDS case definition. European Union Concerted Action on Multinational AIDS Scenarios. *AIDS*, **14,** 2179-2189.

Drake, C. (1993) Effects of Misspecification of the Propensity Score on Estimators of Treatment Effect. *Biometrics*, **49,** 1231.

Dunn, D. T., Newell, M. L., Ades, A. E. and Peckham, C. S. (1992) Risk of human immunodeficiency virus type 1 transmission through breastfeeding. *Lancet*, **340,** 585-588.

ECS and Swiss Cohort and European Collaborative Study and the Swiss HIV Pregnancy Cohort (1997) Immunological markers in HIV-infected pregnant women. *AIDS*, **11,** 1859-1865.

ECS and Swiss Cohort and European Collaborative Study and the Swiss HIV Pregnancy Cohort (2000) Combination antiretroviral therapy and duration of pregnancy. *AIDS*, **14,** 2913-2920.

Ekouevi, D., Inwoley, A., Tonwe-Gold, B., Danel, C., Becquet, R., N'dri-Yoman, T., ROuet, F., Leroy, V., Anglaret, X. and Dabis., F. Criteria for HAART Should Be Revisited in HIV-infected Pregnant Women in Resource-limited Settings. Poster at 13th Conference on Retroviruses and Opportunistic Infections (CROI 2006). Abstract 704a .

EuroHIV . HIV / AIDS Surveillance in Europe: end year report 2003, No. 70. Institut de veille sanitaire, Saint-Maurice, France.

EuroHIV . HIV/AIDS Surveillance in Europe: mid-year report 2005, No. 72. 72. 2005. Saint-Maurice, France, Institut de Veille Sanitaire.

EuroHIV . HIV / AIDS Surveillance in Europe: end year report 2005, No. 73. Institut de veille sanitaire, Saint-Maurice, France. 2006.

European Collaborative Study (1988) Mother-to-child transmission of HIV infection. *Lancet*, **2,** 1039-1043.

European Collaborative Study (1998) Therapeutic and other interventions to reduce the risk of mother-to-child transmission of HIV-1 in Europe. *Br.J.Obstet.Gynaecol.*, **105,** 704-709.

European Collaborative Study (1999) Maternal viral load and vertical transmission of HIV-1: an important factor but not the only one. *AIDS*, **13,** 1377-1385.

European Collaborative Study (2001) HIV-infected pregnant women and vertical transmission in Europe since 1986. *AIDS*, **15,** 761-770.

European Collaborative Study (2002) Level and pattern of HIV-1-RNA viral load over age: differences between girls and boys? *AIDS*, **16,** 97-104.

European Collaborative Study (2003) Exposure to antiretroviral therapy in utero or early life: the health of uninfected children born to HIV-infected women. *J.Acquir.Immune Defic.Syndr.*, **32,** 380-387.

European Collaborative Study (2004) Increased risk of adverse pregnancy outcomes in HIV-infected women treated with highly active antiretroviral therapy in Europe. *AIDS*, **18,** 2337-2339.

European Collaborative Study (2005a) Increasing likelihood of further live births in HIV-infected women in recent years. *BJOG.*, **112,** 881-888.

European Collaborative Study (2005b) Mother-to-child transmission of HIV infection in the era of highly active antiretroviral therapy. *Clin.Infect.Dis.*, **40,** 458-465.

European Collaborative Study (2006) The mother-to-child HIV transmission epidemic in Europe: evolving in the East and established in the West. *AIDS*, **20,** 1419-1427.

European Mode of Delivery Collaboration (1999) Elective caesarean-section versus vaginal delivery in prevention of vertical HIV-1 transmission: a randomised clinical trial. *Lancet*, **353,** 1035-1039.

Everitt, B. S. (1987) *Introduction to Optimization Methods and their Application in Statistics*. Chapman & Hall, London.

Farrington, C. P. (1996) Interval censored survival data: a generalized linear modelling approach. *Stat.Med.*, **15,** 283-292.

Farrington, C. P. and Gay, N. J. (1999) Interval-censored survival data with informative examination times: parametric models and approximate inference. *Stat.Med.*, **18,** 1235-1248.

Fauci, A. S. (1996) Host factors and the pathogenesis of HIV-induced disease. *Nature*, **384,** 529-534.

Fiore, S., Newell, M. L., Trabattoni, D., Thorne, C., Gray, L., Savasi, V., Tibaldi, C., Ferrazzi, E. and Clerici, M. (2006) Antiretroviral therapy-associated modulation of Th1 and Th2 immune responses in HIV-infected pregnant women. *J.Reprod.Immunol.*, **70,** 143-150.

Fischl, M. A., Richman, D. D., Causey, D. M., Grieco, M. H., Bryson, Y., Mildvan, D., Laskin, O. L., Groopman, J. E., Volberding, P. A., Schooley, R. T. and . (1989) Prolonged zidovudine therapy in patients with AIDS and advanced AIDS-related complex. AZT Collaborative Working Group. *JAMA*, **262,** 2405-2410.

Fischl, M. A., Richman, D. D., Grieco, M. H., Gottlieb, M. S., Volberding, P. A., Laskin, O. L., Leedom, J. M., Groopman, J. E., Mildvan, D., Schooley, R. T. and . (1987) The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. A double-blind, placebo-controlled trial. *N.Engl.J.Med.*, **317,** 185-191.

Fitzgerald, A. P., DeGruttola, V. G. and Vaida, F. (2002) Modelling HIV viral rebound using non-linear mixed effects models. *Stat.Med.*, **21,** 2093-2108.

Food and drug administration . Important drug warning – Zerit and Videx. http://www.fda.gov/medwatch/safety/2001/zerit&videx_letter.htm7 . Accessed 2005.

Frater, A. J., Dunn, D. T., Beardall, A. J., Ariyoshi, K., Clarke, J. R., McClure, M. O. and Weber, J. N. (2002) Comparative response of African HIV-1-infected individuals to highly active antiretroviral therapy. *AIDS*, **16,** 1139-1146.

French, R. and Brocklehurst, P. (1998) The effect of pregnancy on survival in women infected with HIV: a systematic review of the literature and meta-analysis. *Br.J.Obstet.Gynaecol.*, **105,** 827-835.

Friedman, J. H. (1984) A variable span scatterplot smoother. *Laboratory for Computational Statistics, Stanford University Technical Report No.5.*

Gabiano, C., Tovo, P. A., de Martino, M., Galli, L., Giaquinto, C., Loy, A., Schoeller, M. C., Giovannini, M., Ferranti, G., Rancilio, L. and . (1992) Mother-to-child transmission of human immunodeficiency virus type 1: risk of infection and correlates of transmission. *Pediatrics*, **90,** 369-374.

Gallant, A. R. (1987) *Nonlinear statistical models*. Wiley and Sons.

Gazzard, B., Bernard, A. J., Boffito, M., Churchill, D., Edwards, S., Fisher, N., Geretti, A. M., Johnson, M., Leen, C., Peters, B., Pozniak, A., Ross, J., Walsh, J., Wilkins, E. and Youle, M. (2006) British HIV Association (BHIVA) guidelines for the treatment of HIV-infected adults with antiretroviral therapy (2006). *HIV.Med.*, **7,** 487-503.

Geretti, A. M. (2006) HIV-1 subtypes: epidemiology and significance for HIV management. *Curr.Opin.Infect.Dis.*, **19,** 1-7.

Ghani, A. C., Henley, W. E., Donnelly, C. A., Mayer, S. and Anderson, R. M. (2001) Comparison of the effectiveness of non-nucleoside reverse transcriptase inhibitor-containing and protease inhibitor-containing regimens using observational databases. *AIDS*, **15,** 1133-1142.

Giolo, S. R. (2004) Turnbull's nonparametric estimator for interval-censored data. *Technical Report.*, Dept. of Statistics, Federal Univesrity of Paraná, Paraná, Brazil.

Giuliano, M., Guidotti, G., Andreotti, M., Liotta, G., Cusato, M., Cristina Marazzi, M., Loureiro, S., Regazzi, M., Palombi, L. and Vella, S. Triple antiretroviral therapy administered during pregnancy and after delivery significantly reduces breast milk viral load. Poster at 13th Conference on Retroviruses and Opportunistic Infections (CROI 2006). Abstract 727.

Goldstein, H. (1989) Efficient prediction models for adult height. *Auxology 88: Perspectives in the science of growth and development*, Smith-Gordon, London., ed. Tanner, J.M.

Grambsch, P. M. and Therneau, T. M. (1994) Proportional Hazards Tests and Diagnostics Based on Weighted Residuals. *Biometrika*, **81**, 515-526.

Gray, L., Cortina-Borja, M. and Newell, M. L. (2004) Modelling HIV-RNA viral load in vertically infected children. *Stat.Med.*, **23**, 769-781.

Greenberg, M., Jacobziner, H., Pakter, J. and Weisl, B. A. G. (1958) Maternal Mortality in the Epidemic of Asian Influenza, New-York-City, 1957. *American Journal of Obstetrics and Gynecology*, **76**, 897-902.

Griffin, J. T., Fraser, C., Gras, L., de Wolf, F. and Ghani, A. C. (2006) The effect on treatment comparisons of different measurement frequencies in human immunodeficiency virus observational databases. *Am.J.Epidemiol.*, **163**, 676-683.

Gurrin, L. C., Scurrah, K. J. and Hazelton, M. L. (2005) Tutorial in biostatistics: spline smoothing with linear mixed models. *Stat.Med.*, **24**, 3361-3381.

Hamers, F. F. and Downs, A. M. (2004) The changing face of the HIV epidemic in western Europe: what are the implications for public health policies? *Lancet*, **364,** 83-94.

Hammer, S. M., Katzenstein, D. A., Hughes, M. D., Gundacker, H., Schooley, R. T., Haubrich, R. H., Henry, W. K., Lederman, M. M., Phair, J. P., Niu, M., Hirsch, M. S. and Merigan, T. C. (1996) A trial comparing nucleoside monotherapy with combination therapy in HIV-infected adults with CD4 cell counts from 200 to 500 per cubic millimeter. AIDS Clinical Trials Group Study 175 Study Team. *N.Engl.J.Med.*, **335,** 1081-1090.

Hammer, S. M., Saag, M. S., Schechter, M., Montaner, J. S., Schooley, R. T., Jacobsen, D. M., Thompson, M. A., Carpenter, C. C., Fischl, M. A., Gazzard, B. G., Gatell, J. M., Hirsch, M. S., Katzenstein, D. A., Richman, D. D., Vella, S., Yeni, P. G. and Volberding, P. A. (2006) Treatment for adult HIV infection: 2006 recommendations of the International AIDS Society-USA panel. *JAMA*, **296,** 827-843.

Health Protection Agency . A complex picture: HIV and other sexually transmitted infections in the United Kingdom: 2006. Health Protection Agency, London.

Hitti, J., Frenkel, L. M., Stek, A. M., Nachman, S. A., Baker, D., Gonzalez-Garcia, A., Provisor, A., Thorpe, E. M., Paul, M. E., Foca, M., Gandia, J., Huang, S., Wei, L. J., Stevens, L. M., Watts, D. H. and McNamara, J. (2004) Maternal toxicity with continuous nevirapine in pregnancy: results from PACTG 1022. *J.Acquir.Immune Defic.Syndr.*, **36,** 772-776.

Ho, D. D., Neumann, A. U., Perelson, A. S., Chen, W., Leonard, J. M. and Markowitz, M. (1995) Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature*, **373,** 123-126.

Holland, P. W. and Wang, Y. J. (1987) Dependence Function for Continuous Bivariate Densities. *Communications in Statistics-Theory and Methods*, **16,** 863-876.

Hughes, J. P. (1999) Mixed effects models with censored data with application to HIV RNA levels. *Biometrics*, **55,** 625-629.

Hytten, F. (1985) Blood volume changes in normal pregnancy. *Clin.Haematol.*, **14,** 601-612.

Ibrahim, T., Moodley, J. and Doorasamy, T. (2004) Lymphocyte changes in pregnancy: a comparison of the human immunodeficiency virus infected and non-infected women. *J.Obstet.Gynaecol.*, **24,** 498-503.

Ickovics, J. R., Wilson, T. E., Royce, R. A., Minkoff, H. L., Fernandez, M. I., Fox-Tierney, R. and Koenig, L. J. (2002) Prenatal and postpartum zidovudine adherence among pregnant women with HIV: results of a MEMS substudy from the Perinatal Guidelines Evaluation Project. *J.Acquir.Immune Defic.Syndr.*, **30,** 311-315.

International Perinatal HIV Group (2001) Duration of ruptured membranes and vertical transmission of HIV-1: a meta-analysis from 15 prospective cohort studies. *AIDS*, **15,** 357-368.

Ioannidis, J. P., Abrams, E. J., Ammann, A., Bulterys, M., Goedert, J. J., Gray, L., Korber, B. T., Mayaux, M. J., Mofenson, L. M., Newell, M. L., Shapiro, D. E., Teglas, J. P. and Wilfert, C. M. (2001) Perinatal transmission of human immunodeficiency virus type 1 by pregnant women with RNA virus loads <1000 copies/ml. *J.Infect.Dis.*, **183,** 539-545.

Iribarren, J. A., Ramos, J. T., Guerra, L., Coll, O., de Jose, M. I., Domingo, P., Fortuny, C., Miralles, P., Parras, F., Pena, J. M., Rodrigo, C. and Vidal, R. (2001) [Prevention of vertical transmission and treatment of infection caused by the human immunodeficiency virus in the pregnant woman. Recommendations of the Study Group for AIDS, Infectious Diseases, and Clinical Microbiology, the Spanish Pediatric Association, the National AIDS Plan and the Spanish Gynecology and Obstetrics Society]. *Enferm.Infecc.Microbiol.Clin.*, **19,** 314-335.

Italian Register for Human Immunodeficiency Virus Infection in Children (2002) Determinants of mother-to-infant human immunodeficiency virus 1 transmission before and after the introduction of zidovudine prophylaxis. *Arch.Pediatr.Adolesc.Med.*, **156,** 915-921.

Jacqmin-Gadda, H., Thiébaut, R., Chene, G. and Commenges, D. (2000) Analysis of left-censored longitudinal data with application to viral load in HIV infection. *Biostatistics*, **1,** 355-368.

Johnstone, F. D., Thong, K. J., Bird, A. G. and Whitelaw, J. (1994) Lymphocyte subpopulations in early human pregnancy. *Obstet.Gynecol.*, **83,** 941-946.

Jones, M. C. (1996) The local dependence function. *Biometrika*, **83,** 899-904.

Jones, M. C. and Koch, I. (2003) Dependence maps: Local dependence in practice. *Statistics and Computing*, **13,** 241-255.

Jungmann, E. M., Mercey, D., DeRuiter, A., Edwards, S., Donoghue, S., Booth, T., Mohan, D., Lyall, H. and Taylor, G. P. (2001) Is first trimester exposure to the combination of antiretroviral therapy and folate antagonists a risk factor for congenital abnormalities? *Sex Transm.Infect.*, **77,** 441-443.

Kaplan, E. L. and Meier, P. (1958) Nonparametric-Estimation from Incomplete Observations. *Journal of the American Statistical Association*, **53,** 457-481.

Kidd, P. (2003) Th1/Th2 balance: the hypothesis, its limitations, and implications for health and disease. *Altern.Med.Rev.*, **8,** 223-246.

Kirkwood, B. R. and Sterne, J. A. C. (2003) *Medical Statistics*. Blackwell Publishing.

Klein, J. P. and Moeschberger, M. L. (1997) *Survival Analysis*. Springer-Verlag New York, Inc..

Kourtis, A. P., Schmid, C. H., Jamieson, D. J. and Lau, J. (2007) Use of antiretroviral therapy in pregnant HIV-infected women and the risk of premature delivery: a meta-analysis. *AIDS*, **21,** 607-615.

Laine, C., Newschaffer, C. J., Zhang, D., Cosler, L., Hauck, W. W. and Turner, B. J. (2000) Adherence to antiretroviral therapy by pregnant women infected with human immunodeficiency virus: a pharmacy claims-based analysis. *Obstet.Gynecol.*, **95,** 167-173.

Laird, N. M. and Ware, J. H. (1982) Random-effects models for longitudinal data. *Biometrics*, **38,** 963-974.

Landes, M., Thorne, C. and Newell, M.-L. Sexually transmitted infections in HIV-1-infected pregnant women: Prevalence and risk factors in a European population. Poster at XVI International AIDS Conference, 2006.Abstract no. TUPE0279 .

Le Chenadec, J., Mayaux, M. J., Guihenneuc-Jouyaux, C. and Blanche, S. (2003) Perinatal antiretroviral treatment and hematopoiesis in HIV-uninfected infants. *AIDS*, **17,** 2053-2061.

Lepage, P., Van de, P. P., Msellati, P., Hitimana, D. G., Simonon, A., Van Goethem, C., Mukamabano, B., Karita, E., Stevens, A. M., Mathieu, G. and . (1993) Mother-to-child transmission of human immunodeficiency virus type 1 (HIV-1) and its determinants: a cohort study in Kigali, Rwanda. *Am.J.Epidemiol.*, **137,** 589-599.

Leyland, A. H. and Goldstein, H. (2001) *Multilevel Modelling of Health Statistics*. John Wiley & Sons, LTD.

Lindsey, J. C. and Ryan, L. M. (1998) Tutorial in biostatistics methods for interval-censored data. *Stat.Med.*, **17,** 219-238.

Loader, C. (1999) *Local Regression and Likelihood.* Springer-Verlag New York, Inc.

Lorenzi, P., Spicher, V. M., Laubereau, B., Hirschel, B., Kind, C., Rudin, C., Irion, O. and Kaiser, L. (1998) Antiretroviral therapies in pregnancy: maternal, fetal and neonatal effects. Swiss HIV Cohort Study, the Swiss Collaborative HIV and Pregnancy Study, and the Swiss Neonatal HIV Study. *AIDS*, **12,** F241-F247.

Lowe, S. H., Prins, J. M. and Lange, J. M. (2004) Antiretroviral therapy in previously untreated adults infected with the human immunodeficiency virus type I: established and potential determinants of virological outcome. *Neth.J.Med.*, **62,** 424-440.

Lucas, G. M., Cheever, L. W., Chaisson, R. E. and Moore, R. D. (2001) Detrimental effects of continued illicit drug use on the treatment of HIV-1 infection. *J.Acquir.Immune Defic.Syndr.*, **27,** 251-259.

Lucas, G. M., Mullen, B. A., Weidle, P. J., Hader, S., McCaul, M. E. and Moore, R. D. (2006) Directly administered antiretroviral therapy in methadone clinics is associated with improved HIV treatment outcomes, compared with outcomes among concurrent comparison groups. *Clin.Infect.Dis.*, **42,** 1628-1635.

Luzzi, G. A., Peto, T. E. A., Weiss, R. A. and Conlon, C. P. (2003) *HIV and AIDS. In Oxford Textbook of Medicine.* Oxford University Press.

Lyles, C. M., Dorrucci, M., Vlahov, D., Pezzotti, P., Angarano, G., Sinicco, A., Alberici, F., Alcorn, T. M., Vella, S. and Rezza, G. (1999) Longitudinal human immunodeficiency virus type 1 load in the italian seroconversion study: correlates and temporal trends of virus load. *J.Infect.Dis.*, **180,** 1018-1024.

Lynch M. and Walsh B. (1998) *Genetics and Analysis of Quantitative Traits.* Sinauer Assocs Inc., Sunderland, MA.

Lyons, F. E., Coughlan, S., Byrne, C. M., Hopkins, S. M., Hall, W. W. and Mulcahy, F. M. (2005) Emergence of antiretroviral resistance in HIV-positive women receiving combination antiretroviral therapy in pregnancy. *AIDS*, **19,** 63-67.

Manavi, K., McDonald, A. and Al Sharqui, A. (2007) Plasma lopinavir trough levels in a group of pregnant women on lopinavir, ritonavir, zidovudine, and lamivudine. *AIDS*, **21,** 643-645.

Mandelbrot, L., Kermarrec, N., Marcollet, A., Lafanechere, A., Longuet, P., Chosidow, D. and Saada, M. (2003) Case report: nucleoside analogue-induced lactic acidosis in the third trimester of pregnancy. *AIDS*, **17,** 272-273.

Mandelbrot, L., Landreau-Mascaro, A., Rekacewicz, C., Berrebi, A., Benifla, J. L., Burgard, M., Lachassine, E., Barret, B., Chaix, M. L., Bongain, A., Ciraru-Vigneron, N., Crenn-Hebert, C., Delfraissy, J. F., Rouzioux, C., Mayaux, M. J. and Blanche, S. (2001a) Lamivudine-zidovudine combination for prevention of maternal-infant transmission of HIV-1. *JAMA*, **285,** 2083-2093.

Mandelbrot, L., Mayaux, M. J., Bongain, A., Berrebi, A., Moudoub-Jeanpetit, Y., Benifla, J. L., Ciraru-Vigneron, N., Le Chenadec, J., Blanche, S. and Delfraissy, J. F. (1996) Obstetric factors and mother-to-child transmission of human immunodeficiency virus type 1: the French perinatal cohorts. SEROGEST French Pediatric HIV Infection Study Group. *Am.J.Obstet.Gynecol.*, **175,** 661-667.

Mandelbrot, L., Peytavin, G., Firtion, G. and Farinotti, R. (2001b) Maternal-fetal transfer and amniotic fluid accumulation of lamivudine in human immunodeficiency virus-infected pregnant women. *Am.J.Obstet.Gynecol.*, **184,** 153-158.

Manfredi, R. and Calza, L. (2006) Nevirapine versus efavirenz in 742 patients: no link of liver toxicity with female sex, and a baseline CD4 cell count greater than 250 cells/microl. *AIDS*, **20,** 2233-2236.

Manfredi, R., Calza, L. and Chiodo, F. (2005) Prospective, open-label comparative study of liver toxicity in an unselected population of HIV-infected patients treated for the first time with efavirenz or nevirapine. *HIV.Clin.Trials*, **6,** 302-311.

Marcollet, A., Goffinet, F., Firtion, G., Pannier, E., Le Bret, T., Brival, M. L. and Mandelbrot, L. (2002) Differences in postpartum morbidity in women who are infected with the human immunodeficiency virus after elective cesarean delivery, emergency cesarean delivery, or vaginal delivery. *Am.J.Obstet.Gynecol.*, **186,** 784-789.

Matthews, G. V., Sabin, C. A., Mandalia, S., Lampe, F., Phillips, A. N., Nelson, M. R., Bower, M., Johnson, M. A. and Gazzard, B. G. (2002) Virological suppression at 6 months is related to choice of initial regimen in antiretroviral-naive patients: a cohort study. *AIDS*, **16,** 53-61.

Mayaux, M. J., Dussaix, E., Isopet, J., Rekacewicz, C., Mandelbrot, L., Ciraru-Vigneron, N., Allemon, M. C., Chambrin, V., Katlama, C., Delfraissy, J. F. and Puel, J. (1997a) Maternal virus load during pregnancy and mother-to-child transmission of human immunodeficiency virus type 1: the French perinatal cohort studies. SEROGEST Cohort Group. *J.Infect.Dis.*, **175,** 172-175.

Mayaux, M. J., Teglas, J. P. and Blanche, S. (2003) Characteristics of HIV-infected women who do not receive preventive antiretroviral therapy in the French Perinatal Cohort. *J.Acquir.Immune Defic.Syndr.*, **34,** 338-343.

Mayaux, M. J., Teglas, J. P., Mandelbrot, L., Berrebi, A., Gallais, H., Matheron, S., Ciraru-Vigneron, N., Parnet-Mathieu, F., Bongain, A., Rouzioux, C., Delfraissy, J. F. and Blanche, S. (1997b) Acceptability and impact of zidovudine for prevention of mother-to-child human immunodeficiency virus-1 transmission in France. *J.Pediatr.*, **131,** 857-862.

Mazhude, C., Jones, S., Murad, S., Taylor, C. and Easterbrook, P. (2002) Female sex but not ethnicity is a strong predictor of non-nucleoside reverse transcriptase inhibitor-induced rash. *AIDS*, **16,** 1566-1568.

McCullagh, P. and Nelder, J. A. (1983) *Generalized Linear Models*. Chapman and Hall.

Mellors, J. W., Munoz, A., Giorgi, J. V., Margolick, J. B., Tassoni, C. J., Gupta, P., Kingsley, L. A., Todd, J. A., Saah, A. J., Detels, R., Phair, J. P. and Rinaldo, C. R., Jr. (1997) Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection. *Ann.Intern.Med.*, **126,** 946-954.

Melvin, A. J., Burchett, S. K., Watts, D. H., Hitti, J., Hughes, J. P., McLellan, C. L., King, P. D., Johnson, E. J., Williams, B. L., Frenkel, L. M. and Coombs, R. W. (1997) Effect of pregnancy and zidovudine therapy on viral load in HIV-1-infected women. *J.Acquir.Immune Defic.Syndr.Hum.Retrovirol.*, **14,** 232-236.

Mirochnick, M. (2000) Antiretroviral pharmacology in pregnant women and their newborns. *Ann.N.Y.Acad.Sci.*, **918,** 287-297.

Mirochnick, M. and Capparelli, E. (2004) Pharmacokinetics of antiretrovirals in pregnant women. *Clin.Pharmacokinet.*, **43,** 1071-1087.

Mocroft, Horban, Clumeck, Stellbrink, Monforte, A., Zilmer, Kirk, Gatell, Phillips and Lundgren (2006a) Comparison of single and boosted protease inhibitor versus nonnucleoside reverse transcriptase inhibitor-containing cART regimens in antiretroviral-naive patients starting cART after January 1, 2000. *HIV.Clin.Trials*, **7,** 271-284.

Mocroft, A., Gill, M. J., Davidson, W. and Phillips, A. N. (1998) Predictors of a viral response and subsequent virological treatment failure in patients with HIV starting a protease inhibitor. *AIDS*, **12,** 2161-2167.

Mocroft, A., Ledergerber, B., Katlama, C., Kirk, O., Reiss, P., d'Arminio, M. A., Knysz, B., Dietrich, M., Phillips, A. N. and Lundgren, J. D. (2003) Decline in the AIDS and death rates in the EuroSIDA study: an observational study. *Lancet*, **362,** 22-29.

Mocroft, A., Phillips, A. N., Ledergerber, B., Katlama, C., Chiesi, A., Goebel, F. D., Knysz, B., Antunes, F., Reiss, P. and Lundgren, J. D. (2006b) Relationship between antiretrovirals used as part of a cART regimen and CD4 cell count increases in patients with suppressed viremia. *AIDS*, **20,** 1141-1150.

Mocroft, A., Youle, M., Morcinek, J., Sabin, C. A., Gazzard, B., Johnson, M. A. and Phillips, A. N. (1997) Survival after diagnosis of AIDS: a prospective observational study of 2625 patients. Royal Free/Chelsea and Westminster Hospitals Collaborative Group. *BMJ*, **314,** 409-413.

Mofenson, L. M., Lambert, J. S., Stiehm, E. R., Bethel, J., Meyer, W. A., III, Whitehouse, J., Moye, J., Jr., Reichelderfer, P., Harris, D. R., Fowler, M. G., Mathieson, B. J. and Nemo, G. J. (1999) Risk factors for perinatal transmission of human immunodeficiency virus type 1 in women treated with zidovudine. Pediatric AIDS Clinical Trials Group Study 185 Team. *N.Engl.J.Med.*, **341,** 385-393.

Montaner, L. J., Griffin, P. and Gordon, S. (1994) Interleukin-10 inhibits initial reverse transcription of human immunodeficiency virus type 1 and mediates a virostatic latent state in primary blood-derived human macrophages in vitro. *J.Gen.Virol.*, **75 (Pt 12),** 3393-3400.

Moore, D. M., Hogg, R. S., Yip, B., Wood, E., Tyndall, M., Braitstein, P. and Montaner, J. S. (2005) Discordant immunologic and virologic responses to highly active antiretroviral therapy are associated with increased mortality and poor adherence to therapy. *J.Acquir.Immune Defic.Syndr.*, **40,** 288-293.

Mucahy, F., Wallace, E., Woods, S., Hanlon, M., Clarke, S., Bergin, C. and Jackson, J. CD4 Counts in Pregnancy Do not Accurately Reflect the Need for Long-term HAART. Poster at 13th Conference on Retroviruses and Opportunistic Infections (CROI 2006). Abstract 704b.

Munoz, A., Gange, S. J. and Jacobson, L. P. (2000) Distinguishing efficacy, individual effectiveness and population effectiveness of therapies. *AIDS*, **14,** 754-756.

Mwanyumba, F., Gaillard, P., Inion, I., Verhofstede, C., Claeys, P., Chohan, V., Vansteelandt, S., Mandaliya, K., Praet, M. and Temmerman, M. (2002) Placental inflammation and perinatal transmission of HIV-1. *J.Acquir.Immune Defic.Syndr.*, **29,** 262-269.

Naumova, E. N., Must, A. and Laird, N. M. (2001) Tutorial in Biostatistics: Evaluating the impact of 'critical periods' in longitudinal studies of growth using piecewise mixed effects models. *Int.J.Epidemiol.*, **30,** 1332-1341.

Nellen, J. F., Schillevoort, I., Wit, F. W., Bergshoeff, A. S., Godfried, M. H., Boer, K., Lange, J. M., Burger, D. M. and Prins, J. M. (2004) Nelfinavir plasma concentrations are low during pregnancy. *Clin.Infect.Dis.*, **39,** 736-740.

Newell, M. L., Rogers, M., Coll, O., Fiore, S., Floridia, M., Giaquinto, C., Grosch-Worner, I., Guiliano, M., Lindgren, S., Lyall, H., Mandelbrot, L., ., Peckham, C., Rudin, C., Semprini, A. E., Taylor, G., Thorne, C. and Tovo, P. A. (2002) Pregnancy and HIV infection: A european consensus on management. *AIDS*, **16 Suppl 2,** S1-18.

Nightingale, S. L. (1998) From the Food and Drug Administration. *JAMA*, **280,** 1472.

O'Shea, S., Newell, M. L., Dunn, D. T., Garcia-Rodriguez, M. C., Bates, I., Mullen, J., Rostron, T., Corbett, K., Aiyer, S., Butler, K., Smith, R. and Banatvala, J. E. (1998) Maternal viral load, CD4 cell count and vertical transmission of HIV-1. *J.Med.Virol.*, **54,** 113-117.

Odell, P. M., Anderson, K. M. and D'Agostino, R. B. (1992) Maximum likelihood estimation for interval-censored data using a Weibull-based accelerated failure time model. *Biometrics*, **48,** 951-959.

Overton, E. T., Sungkanuparph, S., Nurutdinova, D. and Powderly, W. G. (2005) Antiretroviral resistance among HIV-positive pregnant women who have antiretroviral experience from previous pregnancy. *AIDS*, **19,** 1439-1440.

Ozkan, H., Olgun, N., Sasmaz, E., Abacioglu, H., Okuyan, M. and Cevik, N. (1993) Nutrition, immunity and infections: T lymphocyte subpopulations in protein--energy malnutrition. *J.Trop.Pediatr.*, **39,** 257-260.

Palepu, A., Tyndall, M., Yip, B., O'Shaughnessy, M. V., Hogg, R. S. and Montaner, J. S. (2003) Impaired virologic response to highly active antiretroviral therapy associated with ongoing injection drug use. *J.Acquir.Immune Defic.Syndr.*, **32,** 522-526.

Patel, R. R., Steer, P., Doyle, P., Little, M. P. and Elliott, P. (2004) Does gestation vary by ethnic group? A London-based study of over 122,000 pregnancies with spontaneous onset of labour. *Int.J.Epidemiol.*, **33,** 107-113.

Perrin, L., Kaiser, L. and Yerly, S. (2003) Travel and the spread of HIV-1 genetic variants. *Lancet Infect.Dis.*, **3,** 22-27.

Pezzotti, P., Pappagallo, M., Phillips, A. N., Boros, S., Valdarchi, C., Sinicco, A., Zaccarelli, M. and Rezza, G. (2001) Response to highly active antiretroviral therapy according to duration of HIV infection. *J.Acquir.Immune Defic.Syndr.*, **26,** 473-479.

Phillips, A. N., Grabar, S., Tassie, J. M., Costagliola, D., Lundgren, J. D. and Egger, M. (1999) Use of observational databases to evaluate the effectiveness of antiretroviral therapy for HIV infection: comparison of cohort studies with randomized trials. EuroSIDA, the French Hospital Database on HIV and the Swiss HIV Cohort Study Groups. *AIDS*, **13,** 2075-2082.

Phillips, A. N., Staszewski, S., Weber, R., Kirk, O., Francioli, P., Miller, V., Vernazza, P., Lundgren, J. D. and Ledergerber, B. (2001) HIV viral load response to antiretroviral therapy according to the baseline CD4 cell count and viral load. *JAMA*, **286,** 2560-2567.

Pickard, R. E. (1968) Varicella Pneumonia in Pregnancy. *American Journal of Obstetrics and Gynecology*, **101,** 504-508.

Pinheiro, J. C. and Bates, D. M. (2000) *Mixed-Effects Models in S and S-PLUS.* Springer-Verlag New York, Inc.

Podzamczer, D., Ferrer, E., Consiglio, E., Gatell, J. M., Perez, P., Perez, J. L., Luna, E., Gonzalez, A., Pedrol, E., Lozano, L., Ocana, I., Llibre, J. M., Casiro, A., Aranda, M., Barrufet, P., Martinez-Lacasa, J., Miro, J. M., Badia, X., Casado, A., Lupo, S., Cahn, P., Manos, M. and Estela, J. (2002) A randomized clinical trial comparing nelfinavir or nevirapine associated to zidovudine/lamivudine in HIV-infected naive patients (the Combine Study). *Antivir.Ther.*, **7,** 81-90.

Porter, K., Babiker, A., Bhaskaran, K., Darbyshire, J., Pezzotti, P., Porter, K. and Walker, A. S. (2003) Determinants of survival following HIV-1 seroconversion after the introduction of HAART. *Lancet*, **362,** 1267-1274.

Powderly, W. G., Saag, M. S., Chapman, S., Yu, G., Quart, B. and Clendeninn, N. J. (1999) Predictors of optimal virological response to potent antiretroviral therapy. *AIDS*, **13,** 1873-1880.

Public Health Service Taskforce (2006a) Recommendations for use of antiretroviral drugs in pregnant HIV-1 infected women for maternal health and interventions to reduce perinatal HIV-1 transmission in the United States. *http://aidsinfo.nih.gov/ContentFiles/PerinatalGL.pdf*. Accessed January 2007.

Public Health Service Taskforce (2006b). Safety and Toxicity of Individual Antiretroviral Agents in Pregnancy. Accessed February 2007.

Pugh, M., Robbins, J., Lipsitz, S. and Harrington, D. (1992) Inference in the Cox proportional hazards model with missing covariates. *Department of Biostatistics, Harvard School of Publich Health,Boston*, Technical Report 758Z.

R Development Core Team . R: A language and environment for statistical computing. R Foundation for Statistical Computing.Vienna, Austria, 2006.

Read, J. S., Tuomala, R., Kpamegan, E., Zorrilla, C., Landesman, S., Brown, G., Vajaranant, M., Hammill, H. and Thompson, B. (2001a) Mode of delivery and postpartum morbidity among HIV-infected women: the women and infants transmission study. *J.Acquir.Immune Defic.Syndr.*, **26,** 236-245.

Read, J. S., Tuomala, R., Kpamegan, E., Zorrilla, C., Landesman, S., Brown, G., Vajaranant, M., Hammill, H. and Thompson, B. (2001b) Mode of delivery and postpartum morbidity among HIV-infected women: the women and infants transmission study. *J.Acquir.Immune Defic.Syndr.*, **26,** 236-245.

Repik, A., Richards, K. H. and Clapham, P. R. (2007) The promise of CCR5 antagonists as new therapies for HIV-1. *Curr.Opin.Investig.Drugs*, **8,** 130-139.

Rich, K. C., Siegel, J. N., Jennings, C., Rydman, R. J. and Landay, A. L. (1995) Cd4(+) Lymphocytes in Perinatal Human-Immunodeficiency-Virus (Hiv) Infection - Evidence for Pregnancy-Induced Immune Depression in Uninfected and Hiv-Infected Women. *Journal of Infectious Diseases*, **172,** 1221-1227.

Ronsmans, C., Khlat, M., Kodio, B., Ba, M., De Bernis, L. and Etard, J. (2001) Evidence for a 'healthy pregnant woman effect' in Niakhar, Senegal? *Int.J.Epidemiol.*, **30,** 467-473.

Rosenbaum, P. R. and Rubin, D. B. (1983) The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, **70,** 41-55.

Saada, M., Le Chenadec, J., Berrebi, A., Bongain, A., Delfraissy, J. F., Mayaux, M. J. and Meyer, L. (2000) Pregnancy and progression to AIDS: results of the French prospective cohorts. SEROGEST and SEROCO Study Groups. *AIDS*, **14,** 2355-2360.

Saag, M. S., Holodniy, M., Kuritzkes, D. R., O'Brien, W. A., Coombs, R., Poscher, M. E., Jacobsen, D. M., Shaw, G. M., Richman, D. D. and Volberding, P. A. (1996) HIV viral load markers in clinical practice. *Nat.Med.*, **2,** 625-629.

Sabin, C. A. (2004) Pitfalls of assessing hepatotoxicity in trials and observational cohorts. *Clin.Infect.Dis.*, **38 Suppl 2,** S56-S64.

Sabin, C. A., Devereux, H., Phillips, A. N., Hill, A., Janossy, G., Lee, C. A. and Loveday, C. (2000) Course of viral load throughout HIV-1 infection. *J.Acquir.Immune Defic.Syndr.*, **23,** 172-177.

Sabin, C. A., Hill, T., Lampe, F., Matthias, R., Bhagani, S., Gilson, R., Youle, M. S., Johnson, M. A., Fisher, M., Scullard, G., Easterbrook, P., Gazzard, B. and Phillips, A. N. (2005) Treatment exhaustion of highly active antiretroviral therapy (HAART) among individuals infected with HIV in the United Kingdom: multicentre cohort study. *BMJ*, **330,** 695.

Sabin, C. A. and Phillips, A. N. (2001) Treatment comparisons in HIV infection: the benefits and limitations of observational cohort studies. *J.Antimicrob.Chemother.*, **47,** 371-375.

Sarner, L. and Fakoya, A. (2002) Acute onset lactic acidosis and pancreatitis in the third trimester of pregnancy in HIV-1 positive women taking antiretroviral medication. *Sex Transm.Infect.*, **78,** 58-59.

SAS Institute Inc. (2004) *SAS OnlineDoc® 9.1.3*. SAS Institute Inc., Cary, NC.

Saul, J., Erwin, J., Sabin, C. A., Kulasegaram, R. and Peters, B. S. (2001) The relationships between ethnicity, sex, risk group, and virus load in human immunodeficiency virus type 1 antiretroviral-naive patients. *J.Infect.Dis.*, **183,** 1518-1521.

Schmee, J. and Hahn, G. J. (1979) Simple Method for Regression-Analysis with Censored Data. *Technometrics*, **21,** 417-432.

Schulte, J., Dominguez, K., Sukalac, T., Bohannon, B. and Fowler, M. G. (2007) Declines in Low Birth Weight and Preterm Birth Among Infants Who Were Born to HIV-Infected Women During an Era of Increased Use of Maternal Antiretroviral Drugs: Pediatric Spectrum of HIV Disease, 1989-2004. *Pediatrics*, **119**, e900-e904.

Shapiro, R. L., Thior, I., Gilbert, P. B., Lockman, S., Wester, C., Smeaton, L. M., Stevens, L., Heymann, S. J., Ndung'u, T., Gaseitsiwe, S., Novitsky, V., Makhema, J., Lagakos, S. and Essex, M. (2006) Maternal single-dose nevirapine versus placebo as part of an antiretroviral strategy to prevent mother-to-child HIV transmission in Botswana. *AIDS*, **20,** 1281-1288.

Silverman, B. W. (1986) *Density Estimation.* Chapman and Hall.

Slattery, M. M. and Morrison, J. J. (2002) Preterm delivery. *Lancet*, **360,** 1489-1497.

Spire, B., Duran, S., Souville, M., Leport, C., Raffi, F. and Moatti, J. P. (2002) Adherence to highly active antiretroviral therapies (HAART) in HIV-infected patients: from a predictive to a dynamic approach. *Soc.Sci.Med.*, **54,** 1481-1496.

Stebbing, J. and Bower, M. (2006) Comparative pharmacogenomics of antiretroviral and cytotoxic treatments. *Lancet Oncol.*, **7,** 61-68.

Stek, A. M., Mirochnick, M., Capparelli, E., Best, B. M., Hu, C., Burchett, S. K., Elgie, C., Holland, D. T., Smith, E., Tuomala, R., Cotter, A. and Read, J. S. (2006) Reduced lopinavir exposure during pregnancy. *AIDS*, **20,** 1931-1939.

Sterling, T. R., Vlahov, D., Astemborski, J., Hoover, D. R., Margolick, J. B. and Quinn, T. C. (2001) Initial plasma HIV-1 RNA levels and progression to AIDS in women and men. *N.Engl.J.Med.*, **344,** 720-725.

Stern, J. O., Robinson, P. A., Love, J., Lanes, S., Imperiale, M. S. and Mayers, D. L. (2003) A comprehensive hepatic safety analysis of nevirapine in different populations of HIV infected patients. *J.Acquir.Immune Defic.Syndr.*, **34 Suppl 1,** S21-S33.

Sturmer, T., Schneeweiss, S., Avorn, J. and Glynn, R. J. (2005) Adjusting effect estimates for unmeasured confounding with validation data using propensity score calibration. *Am.J.Epidemiol.*, **162,** 279-289.

Szyld, E. G., Warley, E. M., Freimanis, L., Gonin, R., Cahn, P. E., Calvet, G. A., Duarte, G., Melo, V. H. and Read, J. S. (2006) Maternal antiretroviral drugs during pregnancy and infant low birth weight and preterm birth. *AIDS*, **20,** 2345-2353.

Taha, T. E., Brown, E. R., Hoffman, I. F., Fawzi, W., Read, J. S., Sinkala, M., Martinson, F. E., Kafulafula, G., Msamanga, G., Emel, L., Adeniyi-Jones, S. and Goldenberg, R. (2006) A phase III clinical trial of antibiotics to reduce chorioamnionitis-related perinatal HIV-1 transmission. *AIDS*, **20,** 1313-1321.

The American College of Obstetricians and Gynecologists (2000) ACOG committee opinion 234. *Washington DC, ACOG.*

The Antiretroviral Therapy Cohort Collaboration (2006) Rates of disease progression according to initial highly active antiretroviral therapy regimen: a collaborative analysis of 12 prospective cohort studies. *J.Infect.Dis.*, **194,** 612-622.

The European Collaborative Study (1996) Vertical transmission of HIV-1: maternal immune status and obstetric factors. *AIDS*, **10,** 1675-1681.

The Working Group on Mother-To-Child Transmission of HIV (1995) Rates of mother-to-child transmission of HIV-1 in Africa, America, and Europe: results from 13 perinatal studies. *J.Acquir.Immune Defic.Syndr.Hum.Retrovirol.*, **8,** 506-510.

Therneau, T. M. and Grambsch, T. M. (2000) *Modeling Survival Data.* Springer-Verlag New York, Inc.

Thiébaut, R. and Jacqmin-Gadda, H. (2004) Mixed models for longitudinal left-censored repeated measures. *Comput.Methods Programs Biomed.*, **74,** 255-260.

Thiébaut, R., Jacqmin-Gadda, H., Babiker, A. and Commenges, D. (2005) Joint modelling of bivariate longitudinal data with informative dropout and left-censoring, with application to the evolution of CD4+ cell count and HIV RNA viral load in response to treatment of HIV infection. *Stat.Med.*, **24,** 65-82.

Thiébaut, R., Jacqmin-Gadda, H., Chene, G., Leport, C. and Commenges, D. (2002) Bivariate linear mixed models using SAS proc MIXED. *Comput.Methods Programs Biomed.*, **69,** 249-256.

Thiébaut, R., Jacqmin-Gadda, H., Leport, C., Katlama, C., Costagliola, D., Le, M., V, Morlat, P. and Chene, G. (2003) Bivariate longitudinal model for the analysis of the evolution of HIV RNA and CD4 cell count in HIV infection taking into account left censoring of HIV RNA measures. *J.Biopharm.Stat.*, **13,** 271-282.

Thiébaut, R., Jacqmin-Gadda, H., Walker, S., Sabin, C., Prins, M., Del Amo, J., Porter, K., Dabis, F. and Chene, G. (2006) Determinants of response to first HAART regimen in antiretroviral-naive patients with an estimated time since HIV seroconversion. *HIV.Med.*, **7,** 1-9.

Thomson, M. M., Perez-Alvarez, L. and Najera, R. (2002) Molecular epidemiology of HIV-1 genetic forms and its significance for vaccine development and therapy. *Lancet Infect.Dis.*, **2,** 461-471.

Thorne, C., Fiore, S. and Rudin, C. (2003) Antiretroviral therapy during pregnancy and the risk of an adverse outcome. *N.Engl.J.Med.*, **348,** 471-472.

Thorne, C., Newell, M. L., Dunn, D. and Peckham, C. (1995) The European Collaborative Study: clinical and immunological characteristics of HIV 1-infected pregnant women. *Br.J.Obstet.Gynaecol.*, **102,** 869-875.

Thorpe, L. E., Frederick, M., Pitt, J., Cheng, I., Watts, D. H., Buschur, S., Green, K., Zorrilla, C., Landesman, S. H. and Hershow, R. C. (2004) Effect of hard-drug use on CD4 cell percentage, HIV RNA level, and progression to AIDS-defining class C events among HIV-infected women. *J.Acquir.Immune Defic.Syndr.*, **37,** 1423-1430.

Tincani, A., Balestrieri, G., Faden, D. and DiMario, C. (1991) Systemic lupus erythematosus in pregnancy. *Lancet*, **338,** 756-757.

Touloumi, G., Pantazis, N., Babiker, A. G., Walker, S. A., Katsarou, O., Karafoulidou, A., Hatzakis, A. and Porter, K. (2004) Differences in HIV RNA levels before the initiation of antiretroviral therapy among 1864 individuals with known HIV-1 seroconversion dates. *AIDS*, **18,** 1697-1705.

Townsend, C. L., Cortina-Borja, M., Peckham, C. and Tookey, P. (2007a) Mother-to-child-transmission of HIV in the UK and Ireland, 1990-2004. Poster at 14th Conference on Retroviruses and Opportunistic Infections (CROI 2007).Abstract 761 .

Townsend, C. L., Cortina-Borja, M., Peckham, C. S. and Tookey, P. A. (2007b) Antiretroviral therapy and premature delivery in diagnosed HIV-infected women in the United Kingdom and Ireland. *AIDS*, **21,** 1019-1026.

Townsend, C. L., Tookey, P. A., Cortina-Borja, M. and Peckham, C. S. (2006) Antiretroviral therapy and congenital abnormalities in infants born to HIV-1-infected women in the United Kingdom and Ireland, 1990 to 2003. *J.Acquir.Immune Defic.Syndr.*, **42,** 91-94.

Tuomala, R. E., O'Driscoll, P. T., Bremer, J. W., Jennings, C., Xu, C., Read, J. S., Matzen, E., Landay, A., Zorrilla, C., Blattner, W., Charurat, M. and Anderson, D. J. (2003) Cell-associated genital tract virus and vertical transmission of human immunodeficiency virus type 1 in antiretroviral-experienced women. *J.Infect.Dis.*, **187,** 375-384.

Tuomala, R. E., Shapiro, D. E., Mofenson, L. M., Bryson, Y., Culnane, M., Hughes, M. D., O'Sullivan, M. J., Scott, G., Stek, A. M., Wara, D. and Bulterys, M. (2002) Antiretroviral therapy during pregnancy and the risk of an adverse outcome. *N.Engl.J.Med.*, **346,** 1863-1870.

Tuomala, R. E., Watts, D. H., Li, D., Vajaranant, M., Pitt, J., Hammill, H., Landesman, S., Zorrilla, C. and Thompson, B. (2005) Improved obstetric outcomes and few maternal toxicities are associated with antiretroviral therapy, including highly active antiretroviral therapy during pregnancy. *J.Acquir.Immune Defic.Syndr.*, **38,** 449-473.

Turnbull, B. W. (1976) Empirical Distribution Function with Arbitrarily Grouped, Censored and Truncated Data. *Journal of the Royal Statistical Society Series B-Methodological*, **38,** 290-295.

UNAIDS 2006 Report on the Global AIDS Epidemic.

van Benthem, B. H., Vernazza, P., Coutinho, R. A. and Prins, M. (2005) Impact of pregnancy and menopause on CD4 cell counts. *J.Acquir.Immune Defic.Syndr.*, **38 Suppl 1,** S22.

van Heeswijk, R. P., Khaliq, Y., Gallicano, K. D., Bourbeau, M., Seguin, I., Phillips, E. J. and Cameron, D. W. (2004) The pharmacokinetics of nelfinavir and M8 during pregnancy and post partum. *Clin.Pharmacol.Ther.*, **76,** 588-597.

Vaz, M. J., Barros, S. M., Palacios, R., Senise, J. F., Lunardi, L., Amed, A. M. and Castelo, A. (2007) HIV-infected pregnant women have greater adherence with antiretroviral drugs than non-pregnant women. *Int.J.STD AIDS*, **18,** 28-32.

Venables, W. N. and Ripley, B. D. (1997) *Modern Applied Statistics with S-PLUS.* Springer-Verlag New York, Inc.

Villani, P., Floridia, M., Pirillo, M. F., Cusato, M., Tamburrini, E., Cavaliere, A. F., Guaraldi, G., Vanzini, C., Molinari, A., degli, A. A. and Regazzi, M. (2006) Pharmacokinetics of nelfinavir in HIV-1-infected pregnant and nonpregnant women. *Br.J.Clin.Pharmacol.*, **62,** 309-315.

Walmsley, S., Bernstein, B., King, M., Arribas, J., Beall, G., Ruane, P., Johnson, M., Johnson, D., Lalonde, R., Japour, A., Brun, S. and Sun, E. (2002) Lopinavir-ritonavir versus nelfinavir for the initial treatment of HIV infection. *N.Engl.J.Med.*, **346,** 2039-2046.

Ware, J. H. (1985) Linear-Models for the Analysis of Longitudinal-Studies. *American Statistician*, **39,** 95-101.

Watts, D. H., Balasubramanian, R., Maupin, R. T., Jr., Delke, I., Dorenbaum, A., Fiore, S., Newell, M. L., Delfraissy, J. F., Gelber, R. D., Mofenson, L. M., Culnane, M. and Cunningham, C. K. (2004a) Maternal toxicity and pregnancy complications in human immunodeficiency virus-infected women receiving antiretroviral therapy: PACTG 316. *Am.J.Obstet.Gynecol.*, **190,** 506-516.

Watts, D. H., Covington, D. L., Beckerman, K., Garcia, P., Scheuerle, A., Dominguez, K., Ross, B., Sacks, S., Chavers, S. and Tilson, H. (2004b) Assessing the risk of birth

defects associated with antiretroviral exposure during pregnancy. *Am.J.Obstet.Gynecol.*, **191,** 985-992.

Watts, D. H., Lambert, J., Stiehm, E. R., Harris, D. R., Bethel, J., Mofenson, L., Meyer, W. A., III, Mathieson, B., Fowler, M. G. and Nemo, G. (2003) Progression of HIV disease among women following delivery. *J.Acquir.Immune Defic.Syndr.*, **33,** 585-593.

Wegmann, T. G., Lin, H., Guilbert, L. and Mosmann, T. R. (1993) Bidirectional Cytokine Interactions in the Maternal-Fetal Relationship - Is Successful Pregnancy A Th2 Phenomenon. *Immunology Today*, **14,** 353-356.

Weinberg, E. D. (1984) Pregnancy-associated depression of cell-mediated immunity. *Rev.Infect.Dis.*, **6,** 814-831.

Weisser, M., Rudin, C., Battegay, M., Pfluger, D., Kully, C. and Egger, M. (1998) Does pregnancy influence the course of HIV infection? Evidence from two large Swiss cohort studies. *J.Acquir.Immune Defic.Syndr.Hum.Retrovirol.*, **17,** 404-410.

Wimalasundera, R. C., Larbalestier, N., Smith, J. H., de Ruiter, A., McG Thom, S. A., Hughes, A. D., Poulter, N., Regan, L. and Taylor, G. P. (2002) Pre-eclampsia, antiretroviral therapy, and immune reconstitution. *Lancet*, **360,** 1152-1154.

Wood, E., Hogg, R. S., Heath, K. V., de la, R. R., Lee, N., Yip, B., O'Shaughnessy, M. V. and Montaner, J. S. (2003) Provider bias in the selection of non-nucleoside reverse transcriptase inhibitor and protease inhibitor-based highly active antiretroviral therapy and HIV treatment outcomes in observational studies. *AIDS*, **17,** 2629-2634.

Wood, E., Hogg, R. S., Yip, B., Harrigan, P. R., O'Shaughnessy, M. V. and Montaner, J. S. (2004) The impact of adherence on CD4 cell count responses among HIV-infected patients. *J.Acquir.Immune Defic.Syndr.*, **35,** 261-268.

World Health Organisation (2004) HIV transmission through breastfeeding. A review of available evidence. *WHO, Geneva, Switzerland.*.

World Health Organisation (2006) Antiretroviral therapy for HIV infection in adults and adolescents in resource-limited settings: towards universal access. *WHO, Geneva, Switzerland.*.

Yamashita, T. E., Modjarrad, K. and Munoz, A. (2003) Variability of HIV-1 RNA before AIDS and highly active antiretroviral therapy. *AIDS*, **17**, 1264-1266.

Ziegler, J. B., Cooper, D. A., Johnson, R. O. and Gold, J. (1985) Postnatal transmission of AIDS-associated retrovirus from mother to infant. *Lancet*, **1**, 896-898.

Appendix A European Collaborative Study data collection forms

MATERNAL INFORMATION AT DELIVERY

MATERNAL LABORATORY INVESTIGATIONS DURING PREGNANCY AND AT DELIVERY

PERINATAL INFORMATION

ECS3
INTENSIVE PROSPECTIVE STUDY OF CHILDREN BORN TO HIV POSITIVE MOTHERS

MATERNAL INFORMATION AT DELIVERY

Centre
Mothers Study Number
Child Study Number

Mother's date of birth (day, month, year)
Country of birth .....

Marital Status

Single (1), Married (2), Divorced, Separated, Widowed (3), Cohabiting (4)

Ethnic Group

Asian (1), White (2), Black (3), Oriental (4), Other (5)
Age when leaving full-time education, years .....

Obstetric History

Number of previous livebirths .....
Number of previous stillbirths .....
Number of previous miscarriages .....
Number of previous terminations .....

Mothers Risk Group

History of intravenous Drug Abuse (Y/N)
Trimester of last use: pre-conception (0), 1st (1), 2nd (2), 3rd (3), unknown (9)
Needle sharing? never (1) past (2) present (3) unknown (9)
Sexual partner of Bisexual (Y/N)
Sexual partner of Haemophiliac (Y/N)
Sexual partner of Intravenous Drug Abuser (Y/N)
Sexual partner of Other high risk group (Y/N)
(Specify) .....
Other .....

Mothers HIV History

Date of first HIV+ test (day, month, year)

Current clinical status

Current HIV staging (CDC) .....
Specify symptoms .....
Date of onset .....

Details of treatment during pregnancy

Has the woman received any antiretroviral therapy at any time during this pregnancy? Y/N
Please give details of both ART and other prophylaxis (eg. TMP-SMX)

Drug	Date started	Date stopped	Currently taken? (yes/no)

ECS 3
PROSPECTIVE STUDY OF CHILDREN BORN TO HIV POSITIVE MOTHERS

Page 2

MATERNAL INFORMATION

Laboratory investigations during pregnancy and at delivery:

Centre Number 1-2
Mothers Study Number 3-5
Child Study Number 6

Virology

	Date:	Date:	Date:
HIV-**DNA** PCR	Pos / Neg	Pos / Neg	Pos / Neg

HIV-**RNA** PCR	copies/ml	copies/ml	copies/ml
Sample type	Plasma / Serum	Plasma / Serum	Plasma / Serum
Assay used			

Other laboratory investigations

	Date:	Date:	Date:
Total lymphocytes			
CD4 (10^9/litre)			
CD8 (10^9/litre)			
IgG (gm/litre)			
IgA (gm/litre)			
IgM (gm/litre)			
p24 Ag			
HIV Elisa			

ECS3
INTENSIVE PROSPECTIVE STUDY OF CHILDREN BORN TO HIV POSITIVE MOTHERS

PERINATAL INFORMATION

Centre
Mothers Study Number
Child Study Number
Child's date of birth (day, month, year)
Sex (M, F)
Gestational age (weeks)
Birthweight (gm)
OFC (cm

Hospital where delivery took place .....
Obstetrician (initals) .....
Antiretroviral therapy during labour/delivery Y/N
If yes, which drug? .....Orally / IV?
Delivery
Caesarean Section: Elective (1), Emergency (2)
If Caesarean Section, reason .....
Vaginal: Spontaneous (3), vacuum (4), forceps (5)
Presentation: breech (Y/N)
Duration of labour 1st stage (if known) .....
Duration of labour 2nd stage (if known) .....
Time from rupture of membranes to delivery (if known) .....
Scalp Electrodes (Y/N)
Episiotomy or vulvovaginal tear (Y/N)
Perinatal Problems (Y/N). Specify Details:
Hepatomegaly .....
Splenomegaly .....
Drug Withdrawal Symptoms .....
Thrombocytopenic Purpura .....
Infection: suspected (1) confirmed (2) .....
Transfusion ..... *
Congenital Abnormalities ..... *
Other .....
Disposition
with parents (1) fostered (2) adopted (3)
remained in hospital (4) other (5) ..... *
if remained in hospital, say why: .....
Feeding: breast (1) bottle (2) breast and bottle (3)
was breast feeding tried and abandoned? Y/N
Died? Y/N
Date of death: (day/month/year)
Postmortem results, if available ..... *
.....

Take sample required; Please record laboratory results on yellow form

9 2

Appendix B European Collaborative Study Collaborators

Dr C Giaquinto, Dr O Rampon, Dr R D'Elia and Prof A De Rossi (Universita degli Studi di Padova, Italy); Prof I Grosch-Wörner (Charite Virchow-Klinikum, Berlin, Germany); Dr J Mok (Royal Hospital for Sick Children, Edinburgh); Dr I de José, Dra Laarú, Dr I Bates, Dra Salas, Dr J Mª Peña, Dr J Gonzalez Garcia, Dr JR Arribas Lopez and Dr MC Garcia-Rodriguez (Hospital Infantil La Paz, Madrid); Prof F Asensi-Botet, Dr MC Otero, Dr D Pérez-Tamarit (Hospital La Fe, Valencia, Spain); Dr H J Scherpbier, M Kreyenbroek, Dr MH Godfried, Dr FJB Nellen and Dr K Boer (Academisch Medisch Centrum, Amsterdam, The Netherlands); Dr AB Bohlin, Dr S Lindgren, Dr B Anzén, Dr K Lidman, Dr K Elfgren, Dr K Gyllensten and Dr PO Pehrson (Karolinska University Hospital, Huddinge and Solna, Sweden); Prof J Levy, Dr P Barlow, Dr Y Manigart, Dr M Hainaut, Dr A Peltier and Dr T Goetghebuer (Hospital St Pierre, Brussels, Belgium); Dr A Ferrazin (Infectious Diseases Clinic, University of Genoa, Italy); Prof A De Maria (Department of Internal Medicine, University of Genoa, Italy) Prof G Bentivoglio, Dr S Ferrero, Dr C Gotta (Department of Obstetrics and Gynecology-Neonatology Unit, University of Genoa, Italy); Prof A Mûr, Dr A Payà, Dr MA López-Vilchez, Dr R Carreras (Hospital del Mar, Universidad Autonoma, Barcelona, Spain); Dr N H Valerius, Dr V Rosenfeldt (Hvidovre Hospital, Denmark); Dr J Jimenez (Hospital 12 De Octubre, Madrid, Spain); Dr O Coll, Dr A Suy and Dr J M Perez (Hospital Clínic, Barcelona, Spain); Dr C Fortuny, Dr J Boguña (Hospital Sant Joan de Deu, Barcelona, Spain); Dr M Casellas Caro (Hospital Vall D'Hebron, Barcelona, Spain); Dr Y Canet (Hospital Parc Tauli de Sabadell, Barcelona, Spain); Prof G Pardi, Dr M Ravizza (Ospedale San Paolo, Milano, Italy); Dr B Guerra, Dr M Lanari, Dr S Bianchi, Dr L Bovicelli (Policlinico S Orsola, Bologna, Italy); Dr E

Prati, Prof. M. Duse (Universita di Brescia, Brescia, Italy); Dr G Scaravelli, Dr M Stegagno (Universita La Sapienza, Roma, Italy); Dr M De Santis (Universita Cattolica, Roma, Italy); Dr V Savasi, Dr S Fiore, Dr M Crivelli, Prof E Ferrazzi (Ospedale L. Sacco, Milan, Italy); Dr A Viganò, Dr V Giacomet, Dr D Frasca and Prof G Zuccotti (Department of Pediatrics, L. Sacco Hospital, University of Milan); Dr F Ravagni Probizer, Prof A Maccabruni (Policlinico S Matteo, Pavia, Italy); Dr A Bucceri, Dr L Rancilio (Clinica Mangiagalli and Clinica De Marchi, Milano, Italy); Dr S.Alberico, Dr M Rabusin, M Bernardon (IRCCS Burlo Garofolo, Trieste, Italy); Dr G P Taylor, Dr EGH Lyall (St Mary's Hospital, London); Ms Z Penn (Chelsea and Westminster Hospital, London); Drssa W. Buffolano, Dr R Tiseo, (Pediatric Dept, Federico II University, Naples), Prof P Martinelli, Drssa M Sansone, Dr G Maruotti, Dr A Agangi (Obstetric Dept, Federico II University, Naples, Italy); Dr C Tibaldi, Dr S Marini, Dr G Masuelli, Prof C Benedetto (University di Torino, Italy); Dr T Niemieç (National Research Institute of Mother & Child, Warsaw, Poland), Prof M Marczynska, Dr S Dobosz, Dr J Popielska, Dr A Oldakowska (Medical University of Warsaw, Infectious Diseases Hospital, Warsaw, Poland); Dr R Malyuta, Dr I Semenenko, T Pilipenko (ECS Ukraine coordinating centre), Dr S Posokhova, Dr T Kaleeva, (Odessa), Dr A Stelmah, Dr G Kiseleva (Simferopol).

Appendix C Publications arising from this research

Chapter 4:

European Collaborative Study. prepared by Patel, D., Cortina-Borja, M., Thorne, C., Newell ML. Time to undetectable viral load after HAART initiation in HIV-1 infected pregnant women. *Clinical Infectious Diseases, 2007;*44:1647-56

Chapter 6:

European Collaborative Study. prepared by Patel, D., Thorne, C., Fiore, S., Newell, ML. Does Highly Active Antiretroviral Therapy Increase the Risk of Congenital Abnormalities in HIV-Infected Women?. *J.Acquir.Immune Defic.Syndr*; **40**; 116-118

Appendix D R code for nonparametric interval-censored survival estimation (Giolo 2004)

```
cria.tau <- function(data){
l <- data$left
r <- data$right
tau <- sort(unique(c(l,r[is.finite(r)])))
return(tau)
}
S.ini <- function(tau){
m<-length(tau)
ekm<-survfit(Surv(tau[1:m-1],rep(1,m-1)))
So<-c(1,ekm$surv)
p <- -diff(So)
return(p)
}
cria.A <- function(data,tau){
tau12 <- cbind(tau[-length(tau)],tau[-1])
interv <- function(x,inf,sup) ifelse(x(Wimalasundera et al. 2002)>=inf
& x(Mandelbrot et al. 2001a)<=sup,1,0)
A <- apply(tau12,1,interv,inf=data$left,sup=data$right)
id.lin.zero <- which(apply(A==0, 1, all))
if(length(id.lin.zero)>0) A <- A[-id.lin.zero, ]
return(A)
}
Turnbull <- function(p, A, data, eps=1e-3, iter.max=200,
verbose=FALSE){
n<-nrow(A)
m<-ncol(A)
Q<-matrix(1,m)
iter <- 0
repeat {
iter <- iter + 1
diff<- (Q-p)
maxdiff<-max(abs(as.vector(diff)))
if (verbose)
print(maxdiff)
if (maxdiff<eps | iter>=iter.max)
break
Q<-p
C<-A%*%p
p<-p*((t(A)%*%(1/C))/n)
}
cat("Iterations = ", iter,"\n")
cat("Max difference = ", maxdiff,"\n")
cat("Convergence criteria: Max difference < 1e-3","\n")
dimnames(p)<-list(NULL,c("P Estimate"))
surv<-round(c(1,1-cumsum(p)),digits=5)
right <- data$right
if(any(!(is.finite(right)))){
t <- max(right[is.finite(right)])
return(list(time=tau[tau<t],surv=surv[tau<t]))
}
else
return(list(time=tau,surv=surv))
}
```

Appendix E Assessing bivariate relationships

The dependence between two measures such as HIV RNA viral load and CD4 counts can be explored with a scalar dependence such as a correlation, but may not be adequate in summarising the complex structure between the two.

An important assumption made in a bivariate LME model is that the within-group residuals from each response variable are independent. Assessment of the residuals with a plot is a useful tool in assessing any dependence, and may be further aided with a test of statistical significance. The Pearson correlation coefficient, and other scalar dependence measures, used to assess this relationship may be unreliable as the dependence will be averaged out over the range of the residuals. The local dependence map (Jones and Koch 2003), an extension of the local dependence function (Holland and Wang 1987) may be a useful alternative in exploring the dependence between viral load and CD4 and bivariate residuals, which may exhibit non-constant dependence.

Statistical methods

Local dependence function and local dependence maps

The local dependence function (LDF) (Holland and Wang 1987) allows for a more complex understanding of the nature of the dependence between two random variables, X and Y. The LDF is the mixed partial derivative of the log density function and extends the scalar association measure such as the correlation coefficient to one of local dependence (Jones 1996).

Take f to be the density function, such that

$$f^{ij}(x,y)=\frac{\delta f(x,y)}{\delta x^{i}\delta y^{j}}.$$

Consider now discretising the plane by overlaying it with a grid to produce a contingency table with ordered marginal categories, so any point in the grid is now surrounded by a 2 × 2 table made up of the four cells nearest to it. The natural measure of dependence in this local 2 × 2 table is its log-odds ratio. Assuming that the mixed partial derivative of $f(x,y)$ exists and that f is defined on a Cartesian product set, Holland and Wang showed the natural continuous analogue of the collection of the log-odds ratios which characterise the dependency structure of this bivariate discrete data (Jones 1996) to be

$$\gamma(x,y)=\frac{\partial^2 \log f(x,y)}{\partial x \partial y} \quad (1).$$

Jones showed that the global measure of dependence, the Pearson correlation coefficient ρ, could be 'localised' by introducing a kernel function into any integrals involved with this definition and letting the bandwidths of the kernel tend to zero, also resulting in $\gamma(x,y)$ (Jones 1996;Jones and Koch 2003). As a result of these derivations, positive and negative values of γ corresponding to positive and negative dependence can be interpreted in the same way as positive and negative values of the correlation coefficient and log odds ratio. Important properties of the LDF include:

a) it is finite everywhere

b) it is zero if and only if X and Y are independent

c) it is constant if f is the bivariate normal density, and takes the value $\rho/(1-\rho^2)$ elsewhere, where ρ is the Pearson correlation coefficient.

γ can then be estimated with kernel methods using

$$\hat{\gamma}(x,y) = \frac{\hat{f}_{XY}(x,y) - \hat{f}_1(x,y)^{-1}\hat{f}_X(x,y)\hat{f}_Y(x,y)}{(h_1 h_2 k_2)^2 \hat{f}_1(x,y)} \quad (2),$$

where,

$$\hat{f}_Z(x,y) = n^{-1}\sum_{i=1}^{n} Z_i K_{h_1}(X_i - x) K_{h_2}(Y_i - y) \quad (3),$$

$\hat{f}_{XY}$ in (2) is equation (3) with $X_i Y_i$ replacing Z_i, while $\hat{f}_1$ is $\hat{f}_{XY}$ with 1 replacing Z_i and is the usual kernel estimator of f, $\{(X_i, Y_i), i = 1,...,n\}$ is the dataset, h_1 and h_2 are bandwidths associated with the kernel controlling the amount of smoothing in the x and y directions, respectively. The bivariate kernel K is taken to be the product of symmetric univariate densities, where $K_h(u) = h^{-1}K(h^{-1}u)$ is a kernel function with bandwidth h and $k_2 = \int u^2 K(u)du$, is the variance of the kernel function since it has mean zero (Silverman 1986).

The LDF can be graphed using a contour plot, however the results of the plots are hard to interpret, and was the motivation behind developing the local dependence maps (LDM) (Jones and Koch 2003). The construction of the local dependence map can be summarised as follows. The first step is to estimate the density $\hat{f}$ using the kernel density estimator defined in (3), as a product of biweight univariate densities

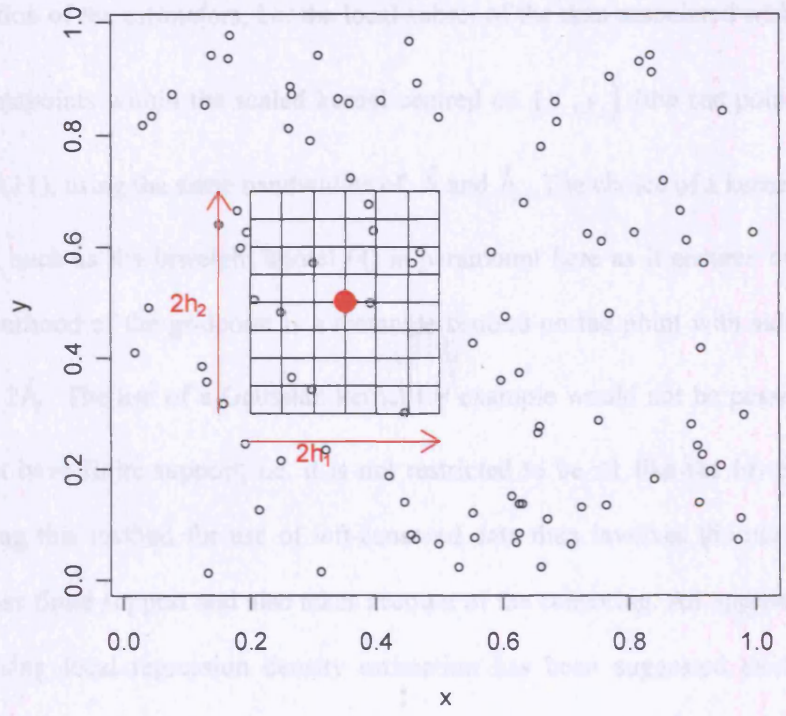
$$K_B(u) = \left(15/16\right)\left(1-u^2\right)^2, \quad u^2 \le 1, \quad (4)$$

with the bandwidths $\hat{h}_1$ and $\hat{h}_2$ chosen as

$$h_i = \frac{s_i}{n^{1/6}} \left\{ \frac{2\sqrt{\pi} \int K^2(u)du}{\int u^2 K(u)du} \right\}^{1/3} \frac{\left(1-r^2\right)^{5/12}}{\left(1+r^2/2\right)^{1/6}}, \; i = 1,2,$$

where s_i is the sample standard deviation and r is the sample correlation. Point at which f appear to be small are then excluded from the LDM, as the estimation of the second derivative, γ will be extremely unreliable there (Jones and Koch 2003). The next step is to carry out a local permutation test of the null hypothesis that $\gamma\left(x_j, y_j\right) = 0$, i.e. a test of local independence. This step reveals the importance of using the biweight kernel over a Gaussian kernel and also highlights why the LDM cannot be extended to the left-censoring case. A regular grid is now placed across each gridpoint, for example at $\left(x_j, y_j\right)$ say, as depicted by the red point shown in Figure A6.1.

Figure A6.1 Example of gridpoint used in permuation testing in LDMs



At this gridpoint, $\hat{\gamma}(x_j, y_j)$ can be calculated using equation (2) and the bandwidths $\hat{h}_1$ and $\hat{h}_2$. Using the data within the grid point, samples satisfying the null hypothesis can be generated by randomly permuting the Y's to break their ties with the X's (with X fixed). This is done P times and $\hat{\gamma}_p$ calculated for each permuted dataset p=1,..., P. Now the local dependence map value is defined to be "positive" if the observed value of $\hat{\gamma}(x_j, y_j)$ is in the highest $(\alpha/2)\%$ of the simulated $\hat{\gamma}_p$'s, "negative" if the observed value is in the lowest $(\alpha/2)\%$ of simulated $\hat{\gamma}_p$'s and set to be "0" otherwise. The recommended value of P=499 was used here with the tests carried out at the 5% significance level $(\alpha = 0.025)$.

Localisation of the permutation test is crucial and is performed in the same way as localisation of the estimators, i.e. the local subset of the data associated with (x_j, y_j) is those datapoints within the scaled kernel centred on (x_j, y_j) (the red point marked in Figure 6.11), using the same bandwidths of $\hat{h}_1$ and $\hat{h}_2$. The choice of a kernel with finite support, such as the biweight kernel (4) is paramount here as it ensures that the local neighbourhood of the gridpoint is a rectangle centred on the point with sides of length $2\hat{h}_1$ and $2\hat{h}_2$. The use of a Gaussian kernel for example would not be possible here, as does not have finite support, i.e. it is not restricted to be ≤ 1 like the biweight kernel. Extending this method for use of left-censored data then involves the use of a kernel which has finite support and also takes account of the censoring. An approach to fit the LDM using local regression density estimation has been suggested using a tricube kernel which also has finite support (Abberger 2002;Loader 1999). Local regression density estimation methods have been extended to include censoring, but do not have the properties of finite support and therefore this approach was not suitable here (Loader 1999;Schmee and Hahn 1979;Buckley and James 1979).

Implementation

The local dependence maps were implemented in R 2.4.1 (R Development Core Team 2006) with the use of a program written by Mario Cortina-Borja (Personal communication).